Influence of syllable-coda voicing on the acoustic properties of syllable-onset /l/ in English

Sarah Hawkins, Noël Nguyen

Influence of syllable-coda voicing on the acoustic properties of syllable-onset /l/ in English

Sarah Hawkins

Department of Linguistics, University of Cambridge, Cambridge, CB3 9DA, UK

Noël Nguyen

Laboratoire Parole et Langage, Université de Provence, 29 avenue Robert Schuman, 13621 Aix-en-Provence, France

PHON 00/021, Journal of Phonetics, revised 31 October 2002.

Running title: Syllable-onset /l/ varies with coda voicing

ABSTRACT

Properties of syllable onset /l/ that depend on the voicing of the syllable coda were measured for four speakers, representing different non-rhotic British English accents that differ in their phonetic realisation of onset /l/ and in their system of phonological contrast involving onset /l/ and /r/. Onset /l/ was longer before voiced than voiceless codas for all four speakers, and darker for two of them as measured by lower F_2 frequency, and for these two and one other as measured by spectral centre of gravity (COG). There were no coda-dependent differences in f_0 in the /l/, and F_1 frequency differed only for the fourth speaker. The vowel was also longer for all four speakers when the coda was voiced (as expected), while F_1 was lower and F_2 normally higher. One speaker provided data with fricative or affricate onsets: fricated segments were longer before voiced codas, but no coda-dependent COG differences were found. At least when the onset includes /l/, phonological voicing of the coda seems to be reflected in complex acoustic-phonetic properties distributed across the whole syllable, some properties being localised, others not. We describe these properties as variations in a bright~sombre dimension. In most accents, when the coda is voiceless, the syllable is relatively bright: small proportions of periodic energy which is relatively high-frequency at the syllable edges, and a high proportion of silence or aperiodic energy. When the coda is voiced, the syllable is relatively sombre: a high proportion of periodic energy which is relatively low-frequency at the syllable edges, and relatively small amounts of silence and aperiodic energy. Other accents use other combinations, dependent on the phonetic and phonological properties of liquids in the particular accent. The association of onset darkness and coda voicing does not seem to be ascribable to anticipatory coarticulation of features essential to voicing itself; this observation provides support for nonsegmental models of speech perception in which fine phonetic detail is mapped directly to linguistic structure without reference to phoneme-sized segments.

1 Introduction

Acoustic-phonetic differences associated with the voiced-voiceless distinction for English stops have been extensively researched, partly because they provide a rich test-bed for investigating speech perception, in that many different acoustic properties can cue the distinction. The focus in looking for these multiple properties for stops in coda position is usually exclusively in the closure or burst of the stop itself, and in the preceding vowel duration (e.g. Jones, 1948; House & Fairbanks, 1953; Denes, 1955; Peterson & Lehiste, 1960; Chen, 1970; Lisker, 1986; Raphael, 1972; Hogan & Rozsypal, 1980; Massaro & Cohen, 1983; Fischer & Ohde, 1990; Kingston & Diehl, 1995).

However, about once a decade, someone shows that coda voicing affects earlier parts of the syllable in perceptually salient ways. After thirty years, the time has come to take note. Here are the data. Wolf (1978) and Summers (1988), using simple VC and bVC stimuli respectively, showed that spectral differences in the early part of low vowels allow American listeners to predict the voicing of the coda. Wolf truncated naturally-spoken /æC/ syllables at a number of points, and concluded from listeners' forced-choice labelling of the final stop that "the cues to voicing in final stops appear to spread throughout the syllable." (1978: 299). Acoustic analyses of her stimuli showed that when the coda was voiced, F_1 was lower and F_2 higher over the first 50 ms of the vowel, the difference being about 38 Hz for F_1 and 75 Hz for F_2. Summers (1988) showed that /bVC/ syllables were more likely to be heard as ending in voiced stops when F_1 of /ɑ/ or /æ/ was lower in onset frequency and steady state (onset- and steady-state frequencies covaried). These findings demonstrated the perceptual salience of Summers' (1987) earlier observation that F_1 frequency is lower in the initial and steady-state parts of the vowel when the following consonant is voiced. Kwong & Stevens (1999) showed the opposite pattern for American pronunciations of the words *writer* and *rider*, in which the medial stop is a voiced tap. For instances of these words that were identified at better than chance in a two-alternative forced choice task, F_1 was lower and F_2 was higher throughout the second part of the diphthong /aɪ/ (including the transition) for *writer*—i.e. for the voiceless member of the pair. No frequency differences were found in the first part of the diphthong. Listeners could use these properties to identify otherwise identical synthetic stimuli as *writer* or *rider*. Kwong and Stevens (1999) interpreted their findings as indicating that the tongue body is raised and fronted and thus the pharynx is expanded in *writer* compared with *rider*, and that this is a remnant of an anticipatory articulatory adjustment to inhibit voicing in the medial /t/, even though that /t/ is now realised as a voiced tap. (The reasoning is that if the expansion happens before the closure for the medial consonant, then it is hard to increase the transglottal pressure-drop during the closure by expanding the pharynx yet more.) They further suggested that such differences are unlikely to appear in lax or low vowels, which do not have offglides in English. Since both Wolf and Summers (op. cit.) used lax and/or low vowels, this may explain the difference between the two earlier studies and the later one. Or, the difference between the studies may reflect a systematic difference between consonants in word-final position (in monosyllables), and medial consonants in disyllables. Note, however, that all three studies have found consistent small differences in formant frequencies throughout considerable portions of the vowel preceding a phonologically voiced or voiceless obstruent coda. In a perceptual experiment using only stylized synthetic stimuli, Whalen (1989 Expt. 1) similarly showed that F_1 frequency interacted with vowel duration in determining whether the stimuli were heard as *bet, bed, bat* or *bad.*

Each of these studies confined its focus to the vowel immediately preceding coda obstruents. In the present study, we asked whether similar coda-dependent differences arise in the syllable onset. This was not an idle question. We are investigating the potential of a nonsegmental theory of spoken word recognition, and in consequence we were searching for a perceptually-salient contrast in fine acoustic-phonetic detail that was caused by the phonological class of a non-adjacent phoneme, preferably one that could not be described in terms of the coarticulatory spread of features that are integral to the production of this conditioning phoneme. Well-attested long-domain coarticulatory spread of lip rounding, nasalization, and so on, were therefore excluded. Wolf's and Summers' perceptual data seemed promising. Although their own acoustic measurements did not exclude the possibility that the perceptual cues early in the vowel stem from differences in jaw height that facilitate the production of voiceless vs. voiced stops, two sources of information suggested that the properties of onset /l/ might vary with the voicing of the tautosyllabic coda.

One of these sources of information comprised measurements on one General American and one British English speaker by van Santen, Coleman & Randolph (1992) and Slater & Coleman (1996). They showed that coda-dependent durational differences for these two speakers were not restricted to syllabic nuclei, but might extend to sonorant consonants, including those in syllable onsets. These measurements were complemented by our second source of information, namely impressionistic observations of longer and darker onset /l/s before voiced codas compared with voiceless ones in many regional accents of British English, most obviously those whose onset /l/s are phonetically dark. British English accents vary widely in the phonetic darkness of their onset /l/s:

the 'Standard Southern British' used in many published phonetics experiments has clear onset /l/, as do many other regional British accents, but a significant number are more like US English in having dark onset /l/, and the degree of darkness differs between these accents. Phonologically, all non-rhotic accents of English maintain a clear/dark contrast with onset liquids (/l r/), but whereas most have a relatively clear onset /l/ and dark onset /r/, others have relatively dark /l/ and clear /r/ (Kelly & Local, 1986, 1989; Carter, 1999, 2002; Local, submitted). Note that this comparison is between the perceived relative quality and/or F_2 frequency of /l/ and /r/ (and surrounding segments) when both are in syllable onsets, and not the much better-known contrast between clear and dark /l/ in onsets vs. codas respectively. That is, these observations compare the relative phonetic quality of segments in word pairs like *array/allay, arrive/alive, mirror/miller, berated/belated*, rather than in pairs like *leaf/feel*. Conceivably, then, voiced codas may condition relatively darker onset /l/s in most accents but clearer /l/s in others. If these observations are right, then it would appear that at least some accents of English mark coda voicing across the whole syllable. A further question in this context is whether the differences are only durational, or involve spectral changes too: although onset /l/s before voiced codas can sound subjectively darker, Newton (1996) showed that onset /l/s in synthetic stimuli may be judged as darker when the only physical difference is that they are longer.

These issues have guided us in a series of studies of syllables which contain onset /l/s and that contrast in the voicing of their coda obstruents. Whereas the focus of most of our previous papers describing this work is perception (Hawkins & Nguyen 2000, 2001, in press), the present paper describes the natural speech measurements upon which our perceptual stimuli and hypotheses have been based. It is worth presenting these acoustic measurements in one paper because the perceptual experiments indicate that durational and spectral properties of onset /l/ can influence listeners' lexical decisions. When a syllable-onset /l/ is relatively long and dark, the word is more likely to be heard as having a voiced rather than a voiceless coda. Hawkins & Nguyen (in press) showed this in a lexical-decision task in which the stimuli were cross-spliced onsets and rhymes of naturally-spoken real and nonsense words; Hawkins & Nguyen (2000, 2001) also showed it in a series of experiments involving forced-choice identification of synthetic tokens of *led* and *let* in which only the /l/ and the beginning of the vowel were audible, the later part of the syllable being replaced by noise. Hawkins & Nguyen's (2001) results indicated that this type of information is used by listeners when it is systematic in the heard speech; listeners seem to learn quickly whether it is worth using a particular acoustic property as a cue to a lexical distinction (see also Smith 2001).

The data of the present study come from Hawkins & Nguyen (in press). Their perceptual stimuli were taken from a larger set of real speech that was collected and partially analysed before running the perceptual experiments, in order to test the hypothesis that there would be differences in the duration and possibly the spectrum of onset /l/ dependent upon the voicing of the coda. The present study measures durational and spectral properties of the syllables, for reasons explained above, and also the fundamental frequency, because Hawkins & Nguyen (2000, 2001) found that listeners associated a lower f0 with voiced rather than voiceless codas in their forced-choice task using synthetic stimuli. Some of the control data collected for the lexical decision task (Hawkins & Nguyen in press) comprised FVC syllables in which F was a voiceless fricative or affricate. Durational and spectral measures of the onset frication of those syllables are described in Appendix 2 of this paper.

2 Method

2.1 Material

The experimental material consisted of 39 pairs of (C)lVC monosyllables differing in the voicing of the final stop (voiced for one member of the pair and voiceless for the other). Syllable onsets were /l bl pl gl kl fl sl/; vowels were /i ɪ eɪ eə ɛ æ aɪ ɑ ɒ u ʌ ɜ əʊ/. The words were chosen to meet the requirements of a perceptual experiment that used a lexical decision task, as described in Hawkins and Nguyen (in press). Consequently, half were real English words, and the other half nonsense. For 14 pairs, the voiced member was a real word, e.g. *load,* and the voiceless member was a non-word, *loat*; the reverse was true for the other 25 pairs (i.e. the real words had voiceless codas e.g. *sleet sleed*). To conform to the requirements of the perceptual experiment, and possibly important for a production study that mixes real and nonsense words (*cf.* Wright, in press), all words had frequencies of less than 50 per million (mean 4.41 in the Brown Corpus). We controlled as far as possible the frequencies of the lexical competitors for each pair, defined as other English monosyllables beginning with the same (C)lV sequence. Most pairs (33) only had low-frequency lexical competitors (<100 per million). Three pairs had one high-frequency competitor, and three had two high-frequency competitors. All words are listed in Appendix 1. The words were produced both in citation form and in a carrier phrase. The carrier phrase, *Put up a*

___ *above all*, was designed to minimize potential lingual coarticulatory interactions between the critical monosyllable and the flanking sounds. Nuclear stress was on *all*.

2.2 Speakers

The 4 subjects were native speakers of British English. All except S4 had some phonetic training, and all except S1 were naive as to the purpose of the study. All were teachers or students at Cambridge University. They represent four different non-rhotic regional accents of British English that were chosen because they were predicted to exhibit systematic spectral and durational variations in onset /l/s. S1, the first author, was in her 40s and has lived in several regions of England and the USA. At the time of the recording she had lived in Cambridge, England for the preceding 11 years, and for this experiment she maintained an Educated Northern British accent, with clear onset /l/ and dark onset /r/. S2-S4 were males aged mid-20s to early 30s. S2 is mildly RP with Manchester attributes including dark onset /l/ and even darker onset /r/. S3 has a strong northeastern (York) accent with fairly dark onset /l/ and darker onset /r/. Thus, though they have distinctly different regional accents, S1-S3 all have lighter onset /l/ than onset /r/. In contrast, S4, with a strong regional accent from the north west (St. Helens, Lancashire) has the only accent of the four to realise the contrast in darkness between onset /l/s and /r/s in the opposite way to that of most British English regional accents: a relatively darker onset /l/ than /r/; his onset /l/ is very dark.

Thus the four subjects can be grouped in two different ways. Phonetically, S2, S3, and S4 form a group that is distinct from S1 in having a dark rather than clear onset /l/. That is, the onset /l/ that S2, S3, and S4 use in words like *let* and *led* would be normally be [l^{γ}] in a systematic narrow transcription, whereas S1's would be [l]. Phonologically however, S1, S2, and S3 form a group that is distinct from S4, in that each person's onset /l/ is relatively clearer than his/her onset /r/. In other words, S1 and S4 differ in both their phonetic realisation of /l/, and in the system of phonological contrast that onset /l/ has with onset /r/. S2 and S3 are like S4 in their phonetic realisation of /l/, but like S1 in their phonological system for the relative darkness of onset /l/ and /r/.

2.3 Procedure

The recordings were made in a sound-treated room using high-quality equipment. Each of the 4 speakers read the 78 monosyllables in random order, five times in isolation, and five times in a carrier phrase (recorded in separate blocks), for a total of 3,120 tokens. The material to be read appeared on a computer screen in front of the subject, one word or word-plus-carrier at a time. Speakers were told to speak naturally, but with consistent rate, pitch contour, and stress. As each item was spoken, both experimenters monitored its quality, by listening and by examining its waveform (low-pass filtered, digitized at 16 kHz SR) displayed in real time on a Silicon Graphics workstation. Disfluent or otherwise unsatisfactory items were repeated immediately.

2.4 Segmentation criteria

The data were segmented and labelled by hand from waveforms and wide-band spectrograms. For each of the 3,120 items, we marked the onset of /l/, the start of its periodic part where this was different from its onset (as when it was clustered with a voiceless obstruent, e.g. /sl/), /l/ offset, and vowel offset. Criteria for defining /l/ onset depended on the syllable onset, as follows. Single /l/ preceded by a vowel (i.e. in the carrier): an abrupt fall in overall intensity and formant frequencies. Single /l/ in isolated items: the start of periodicity. Stop+/l/ clusters: the transient at stop release. In /sl/ clusters: where the aperiodic noise abruptly decreased in intensity and its spectral shape changed, sometimes immediately after a high-amplitude transient. In /fl/ clusters: typically, at an abrupt increase in noise intensity and associated spectral changes. For all items, /l/ offset was segmented at the abrupt rise in formant amplitudes and, usually, frequencies. Vowel offset was at the end of periodicity; glottalisation was included as part of the vowel.

2.5 Measurements

Six types of measurement were made, using *xwaves+* and ESPS on the data digitized at 16 kHz SR unless otherwise stated in the text. They were as follows.

(1) Durations of the onset /l/ and of the vowel.

(2) The frequencies of F_1 and F_2 at three points in the syllable: the midpoint of the periodic part of /l/ (l-mid), 40 ms after /l/ offset (V-on, which was roughly at the end of the /l/-to-V transition), and at the midpoint of the vowel (V-mid).

(3) The spectral center of gravity (COG) at five points in the syllable: at the same locations as F_1 and F_2 were measured, l-mid, V-on, and V-mid; and also at l-end, 24.5 ms before /l/ offset (so the right hand edge of the spectral window fell at the /l/-vowel segmentation point), and at V-end, 30 ms before the end of the vowel.

(4) Duration and rise of F_1 and F_2 transitions (on a few stimuli from Speaker 1 only).

(5) The fundamental frequency at the same five locations at which the COG was measured in (3).

(6) In addition to the five "time-warped" measures in (3), an "unwarped" comparison of the COG was made in the vowel by comparing the COGs at the midpoint of the vowel in the voiceless member of the pair, and at the same time temporal location post-vowel onset in the voiced member. This unwarped measure was included in case it proved more meaningful to compare changes in the spectral COG in terms of their time-course rather than their position relative to the beginning and end of the vowel. However, it revealed no insights, and the data are not reported.

To measure the first two formant frequencies the signal was down-sampled to 10 kHz, high-pass filtered above 80 Hz, and pre-emphasized with a constant of 0.7. Then *xwaves'* automatic formant tracker was used (12-pole autocorrelation lpc spectra, and a step size of 10 ms with both 49-ms and 100-ms $\cos^4$ windows). For F_2, the value from the 49-ms window was used as long as the two measures differed by less than 50 Hz and each was lower than 1800 Hz. The 33% of cases which failed these criteria were remeasured manually from dft spectra supplemented by lpc spectra and spectrograms. For F_1, measured frequencies for both windows had to fall between 200 Hz and 1000 Hz; the 22% of cases that failed these criteria were likewise measured by hand.

The spectral centre of gravity offers a global picture of spectral differences across the entire voiced part of the syllable. It has the advantage of giving a single measure and (unlike formant frequencies) can be done fully automatically, without hand-checking for misidentified formants. To calculate the COG (freeware from Zheng *et al.*), a dft spectrum (50-ms Hanning window centred at the specified place on the pre-emphasized signal, pre-emphasis constant 0.94) was converted into an auditory excitation pattern with decibel magnitude (Moore & Glasberg, 1987). The COG was computed from this excitation pattern between 50 Hz and 3500 Hz.

The f_0 was measured using ESPS's automatic tracker with *xwaves'* default settings and a step size of 5 ms.

3 Results

3.1 Preliminary analyses

Lexical status. There were no differences in the duration or F_2 frequency of /l/ due to lexical status, confirming our impression that the non-words were read fluently. The analyses allowing this conclusion were repeated-measures ANOVAs for each speaker with lexical status, ±carrier phrase, and coda voicing as independent variables.

Vowel quality and syllable onset type. A broad distinction was made between four basic vowel categories (high-front, low-front, low-back/central, back-rounded) in the statistical analyses. The vowels were classified into these categories as follows:

high-front	/i ɪ eɪ aɪ/	(10 paired items)
low-front	/ɛ eə æ/	(7 paired items)
low-back/central	/ɜ ɑ (ʌ)/	(11 paired items)
back-rounded	/ɒ əʊ u ʌ/	(11 paired items)

The subjects' accents required that /ʌ/ was placed in the back-rounded category for S2, S3, and S4, and in the low-back/central category for S1. (S2, S3, and S4 pronounced *flood, glug, glut, slut* and their non-word pairs with the same vowel as in *foot*, [ʊ]; their accents have no [ʌ] sound and lack the phoneme normally transcribed as /ʌ/. S1's accent includes both /ʌ/ and /ʊ/ phonemes, with [ɐ] (phoneme /ʌ/) used for *flood, glug, glut,* and *slut*, and [ʊ], phoneme /ʊ/, used for *foot*.)

Each of the measures reported in the sections below was analysed for each speaker separately using three repeated-measures ANOVAs, the vowel ANOVA, the syllable-onset ANOVA, and the pooled ANOVA. In all ANOVAs, pairs of items was a random variable. The independent variables in the vowel ANOVA were context (with or without carrier), vowel category (as shown above), and coda voicing (voiced or voiceless). In the syllable-onset ANOVA, they were context, syllable onset, and coda voicing. Vowel Category and Syllable Onset could not be in the same ANOVA because the constraints imposed on the corpus by the companion perceptual experiment precluded all possible syllable-onset x vowel category combinations. Independent variables in the pooled ANOVA were context and coda voicing. The pattern of results for the main parameter of interest, coda voicing, was identical in all three ANOVAs. Thus, for simplicity, only the results of the pooled ANOVA are reported, unless otherwise specified.

3.2 Duration of onset /l/

Table I shows that, as expected, the mean duration of onset /l/ was longer before voiced codas than before voiceless ones, and there was remarkably little difference between speakers. At only about 4.2 ms, the mean differences were small, but they were strongly significant for all Ss, and (except for a significant interaction with syllable onset for S3) they were independent of all other independent variables in all three ANOVAs. (In the pooled ANOVAs for S1, S2, S3, and S4 respectively, /l/ duration $F(1,38) = 35.82, 20.84, 16.07, 17.42$, $p < 0.001$ in each case.)

TABLE I ABOUT HERE

There were predictable differences due to speaking context (presence vs. absence of carrier phrase) and syllable onset, but no influence of vowel quality on /l/ duration. Thus, /l/ was 3-19 ms longer in isolated monosyllables than in those spoken in the carrier phrase, the difference being significant for three Ss (S1 (8 ms), S3 (19 ms), and S4 (8 ms)) in all three ANOVAs, and in the syllable-onset ANOVA for S2 (3 ms). (Speaking context for S1, S3, and S4 respectively: $F(1,38) = 24.69, 134.91, 20.23$, $p < 0.001$ in each case. For S2, pooled $F(1,38) = 2.14$, $p = 0.152$; syllable-onset $F(1,32) = 8.6$, $p = 0.006$).) These differences due to carrier-phrase presence were independent of coda voicing. The duration of /l/ varied with syllable onset in expected ways ($p < 0.001$ for each S), so that, for example, /l/ as measured was longer in isolation than after a voiceless stop.

3.3 Duration of the vowel

Vowel durations followed the expected pattern of being longer before voiced codas than voiceless codas. Table II shows, for each speaker, the duration of the vowel in voiced and voiceless contexts (columns 2 and 3 respectively), together with the differences between them (column 4), and their ratios (column 5). Differences due to coda voicing are significant at well beyond the 0.0001 level of probability in all ANOVAs. These data are unremarkable, except that it may be worth noting that S3, who has the smallest coda-dependent durational differences in the vowel, has the largest spectral differences in the /l/, as discussed in the next two sections.

TABLE II ABOUT HERE

3.4 Spectral shape: measurements of F_1 and F_2 frequency

FIGURE 1 ABOUT HERE

Figure 1 shows the average differences in F_1 and F_2 frequency depending on coda voicing at the three time locations defined above. The differences are expressed as the change in F_1 or F_2 frequency when the coda is voiced relative to its frequency when the coda is voiceless. Thus, differences are negative when the formant frequency was lower before a voiced coda than before a voiceless one. Statistically significant differences ($p < 0.05$) are shown in black; non-significant differences in grey. Mean frequencies for each speaker are shown in Table III; F ratios and associated *p* values for the significant differences are given in Table IV.

TABLE III ABOUT HERE

TABLE IV ABOUT HERE

Mean differences for all four speakers are in the right panel. At the midpoint of the onset /l/, both F_1 and F_2 were lower when the coda was voiced. Thereafter, the differences diverge in the two formants, so that F_1 becomes increasingly lower before a voiced coda relative to a voiceless one, whereas F_2 becomes increasingly higher before a voiced coda relative to a voiceless one, until, by the midpoint of the vowel, the difference in F_2 is in the opposite direction—slightly (but not significantly) higher before a voiced coda.

The left-hand panels in Figure 1 show that each speaker followed broadly the same pattern as the above description, though differing somewhat in which differences were significant, and, for F_2, in the size and direction of the difference at the midpoint of the /l/. They all had similar patterns of F_1: it was lower in syllables with voiced codas, not always significantly so in the /l/ but always significantly at the two locations in the vowel, and this difference tended to increase in magnitude at the mid-point of the vowel compared to the /l/ and the vowel onset. (The pattern for F_1 is considered in more detail below.) The overall pattern for F_2 was also similar in all four speakers between the /l/ and the vowel mid-point, but the details of how it was achieved differ. Overall, the balance gradually shifted, with F_2 frequency becoming increasingly higher before a voiced coda compared with a voiceless one. Specifically, for S2 and S3 the difference in F_2 was quite negative in the

/l/, while it was close to zero for S1 and S4. For every speaker, however, the difference evolved towards a more positive value in the vicinity of the /l/-V boundary and thereafter to the vowel midpoint. Thus, at any one of the three measurement points, the speakers differed amongst themselves more for F_2 than for F_1, but since for F_2 they all produced the same increasingly-positive trend through the first half of the syllable, the implication is that they differed more in relative phonetic quality than in the general type of coda-dependent difference introduced.

Although the coda-dependent voiced-voiceless difference in F_2 is not significant overall or for any individual speaker at the vowel midpoint, it does reach significance for two speakers if the effects of vowel quality are taken into account. The vowel x coda voicing interaction was significant for S1 and S2 ($F(3,35) = 7.845$, $p < 0.001$; $F(3,35) = 3.823$, $p < 0.02$ respectively), in both cases because the difference was negative for the back-rounded vowel group, yet positive for the other three groups. Figure 2 shows that when the back-rounded group is excluded from the analysis, the coda-dependent difference in F_2 frequency at the midpoint of the vowel reaches significance for S1 and S2 (for S1, $F(1,24) = 17.463$, $p < 0.001$; for S2, $F(1,24) = 9.758$, $p = 0.005$).[1]

FIGURE 2 ABOUT HERE

Figure 2 also shows that the significant lowering of F_1 in onset /l/ before voiced codas in S2's speech is due to the words with back rounded vowels; when /l/ preceded other vowels, there is only a small, nonsignificant negative trend in F_1 at l-mid. In other words, while S2 had a tendency to lower F_1 in the middle of /l/ before voiced codas, the significance of the difference depends on the back rounded vowel group being included in the analysis. S1 continues to show no difference in F_1 at l-mid. Thus only S4 can be considered to normally lower F_1 in onset /l/ before voiced codas.

We conclude that, although the observed coda-dependent differences in F_1 frequency in the /l/ are consistent with the pattern of variation for F_1 over the following vowel, this result should be interpreted with caution. First, the interactions with vowel quality, discussed above with respect to Figure 2, suggest that the mean data in Figure 1's righthand panel do not reflect the most general case: only S4 has a reliably lower F_1 at l-mid. Second, there were methodological difficulties in measuring F_1 frequency. Recall that 22% of the automatically-measured F_1 values in the /l/ were judged to be unreliable and were therefore remeasured by hand. However, in most of these hand-measured tokens, the /l/ had a very short periodic portion (4-6 periods) because it occurred in a consonant cluster with a preceding voiceless obstruent. The spectra used for measuring F_1 were therefore computed over a very short time window, at the expense of accuracy in frequency. In addition, a considerable proportion of the periodic portion of these /l/s (c. 1-3 periods) often comprised only very low-frequency periodicity at voicing onset, with almost no periodic energy much above *f0*; this may have biased the measured F_1 frequency downwards. Wide formant bandwidths during the aspiration in a preceding voiceless stop may also have contributed to introducing inaccuracies in the manual estimation of F_1 frequency. We thus consider these hand measurements to be relatively unreliable, just like the automatic measurements on the same utterances. When we restricted the data to the automatically-detected F_1 values that met our criteria for reliability (i.e. when we excluded hard-to-measure tokens), there were no significant differences in F_1 frequency in /l/ depending on coda voicing for any of the speakers, including S4. On balance, we conclude that there seems usually to be little or no difference in F_1 frequency in the /l/ itself. Arguably the most important thing, however, is that F_1 varied in a consistent way across the /l/+vowel sequence for the four speakers, as indicated above, with a tendency for an increasingly low F_1, as one progresses through the syllable, before a voiced coda as opposed to a voiceless one.

FIGURE 3 ABOUT HERE

Figure 3 summarises the combined changes in F_1 and F_2 frequency for each speaker at the three locations, l-mid, V-on, and V-mid. It plots the same data as in Figure 1, but as a single data point for each speaker at each location, rather than separately for each formant, by showing the mean difference (Hz) between F_2-F_1 in syllables with voiced codas, and F_2-F_1 in syllables with voiceless codas. For example, $[F_2\text{-}F_1]_{load} - [F_2\text{-}F_1]_{loat}$, and so on over all voiced-voiceless pairs. Negative differences thus indicate that the difference between the first two formants was greater when the coda was voiceless; positive differences indicate that the difference between the first two formants was greater when the coda was voiced. Reference to Figures 1 and 2 shows that these differences are normally likely to reflect changes in the relative frequencies of both F_1 and F_2, though usually to different degrees in different parts of the syllable, and to different degrees for different speakers. It seems reasonable to infer that a greater difference between the first two formants in any given syllable can be taken as indicating that the articulation is more fronted and more open, though sometimes the change is greater in one dimension than the other, and this interpretation is more tentative for /l/ than for the vowels. However, adopting it as a broad generalisation, then negative differences in Figure 3 suggest that the articulation is relatively more backed and/or lower before a voiced coda. Positive differences indicate that it is relatively more fronted and/or higher before a voiced coda. All four speakers have positive differences at the two measured points in the vowel. S4 also has a positive difference at l-mid, whereas the other three speakers all have negative differences of different sizes, substantial in the case of S3, and essentially zero in the case of S1.

Generalisation from these data must be cautious because we have only one speaker representing each accent, but nevertheless, because nuances of regional accent *are* carefully distinguished in this study, it is noteworthy that Figure 3 illustrates simply both the phonological and the phonetic differences in onset /l/ that are known to distinguish these accents. Recall that the speakers/accents fall into two groups phonologically, and two different groups phonetically. Phonologically, S1, S2, and S3 form one group, in that they all have clearer onset /l/ than onset /r/; these three speakers all have a negative or zero difference at the midpoint of the /l/ in Figure 3. In contrast, S4 has the only accent in which onset /l/ is darker than onset /r/, and he is unique amongst the four speakers in having a large positive difference at l-mid in Figure 3. In terms of phonetic realisation of the onset /l/, S1 is distinguished from the other two speakers in the same phonological group (S2 and S3) in that the phonetic quality of all S1's onset /l/s is clear, whereas the quality of all S2's and S3's is dark; and S1 produces the same difference between F_1 and F_2 before both voiced and voiceless codas (i.e. essentially no difference at l-mid in Figure 3), whereas S2 and S3 have a backer or lower articulation when onset /l/ is in a syllable with a voiced coda (significant negative difference at l-mid for S3 in Figure 3 ($F(1,35) = 7.717$, $p = 0.009$), and a nonsignificant negative trend for S2 ($F(1,35) = 0.968$, $p = 0.332$). Reference to Figures 1 and 2 shows that S2 and S3 achieve their effect mainly by lowering F_2 rather than F_1, whereas S1 lowers neither. Assuming that lower F_2 without concomitant changes in F_1 reflects tongue dorsum backing and possibly velarisation (Stevens 1998), these data thus suggest that, whereas S2 and S3 appear to further darken their already dark onset /l/ before voiced codas, S1's clear onset /l/ varies little in clarity before voiced and voiceless codas. We return to this point in the next section, on spectral centre of gravity of these syllables, and again in the Discussion.

3.5 Spectral shape: Centre of Gravity

Figure 4 shows the spectral Centre of Gravity (COG) measured at all five locations defined above, expressed as mean differences between voiced-coda and voiceless-coda syllables. Figure 4 thus mirrors Figure 1 except that it shows COG rather than formant frequency, and five locations rather than three; additionally, the right hand panel shows mean COG for only S1, S2, and S3, since S4 had a very different COG pattern in the /l/ as might be expected since he differs from the other speakers in his onset liquid system. The F ratios and the associated *p* values for the significant differences (black rectangles in Figure 4) are given in Table V.

FIGURE 4 ABOUT HERE

Our aim in using the COG was that it can be computed completely automatically, with no hand-checking. This contrasts with the considerable time required to hand measure individual formant frequencies when automatic formant tracking fails to satisfy stringent reliability criteria. There would thus be significant savings in time if the COG could capture the effect of coda voicing on the formant pattern in the onset /l/ and the vowel over an extended time range. However, it was not certain that the COG would be sensitive enough to reflect the relatively small coda-dependent changes in formant frequency that occur early in the syllable.

Figure 4 shows that the COG at the end of the vowel (V-end, 30 ms from the end of the vowel) is lower when the coda is voiced. The difference between voiced and voiceless contexts is strongly significant for S1, S2, and S4; and it is stronger when the item was spoken in isolation than in the carrier phrase for S1 and S2, as indicated by significant coda x carrier interactions for these two speakers (see Table V). This pattern presumably reflects the well-known difference in transitional formant structure at voiced vs. voiceless vowel-obstruent boundaries, with the falling F_1 transition being excited for longer when the coda obstruent is voiced. The pattern confirms that our computed COG does reflect relatively large and consistent differences in formant frequency.

TABLE V ABOUT HERE

The non-significance of the differences at V-on and V-mid confirm that the COG is not a useful measure for coda-dependent differences at earlier points in the vowels, as it obscures the opposite trends in F_1 and F_2 shown in Figures 1 and 2. It is, however, a reasonable reflection of F_2 in onset /l/ for S2, S3, S4—all those with dark onset /l/: the COG is lower in onset /l/ when the coda is voiced for S2 and S3, and unchanged for S4, despite S4's lower F_1 when the coda is voiced. The COG is not a good reflection of either F_1 or F_2 for S1, but on the other hand, it does show that S1 produces lower overall spectral balance in the onset /l/ before voiced codas, like S2 and S3, with whom she shares the same phonological relationship between onset /l/ and /r/. Presumably

S1 lowers frequencies or amplitudes in higher parts of the spectrum, but does not systematically lower F_2 in onset /l/. She might perform a similar manoeuvre throughout the syllable, since she is unique in producing the entire (C)lVC syllable with lower COG before voiced codas. Although S1's lower COG in later parts of the vowel could be due to the significant difference in F_1, this argument is tenuous since Figure 1 shows that S1's patterns for F_1 and F_2 in the vowel are similar to those of the other three speakers. In other words, whereas S2 and S3 confine their largest and most consistent spectral difference to the onset /l/, S1 seems to maintain a smaller but more consistent difference throughout the syllable. The implication is that S1 exploits some other way of lowering spectral frequencies in syllables that have voiced rather than voiceless codas. This distinctive pattern merits more investigation with a larger group of speakers, because S1 is the only speaker amongst the four with a clear onset /l/.

In conclusion, these data suggest that the COG offers some potential for tracking small coda-dependent spectral changes in onset /l/ itself, but not in the first half of the vowel. Its value for onset /l/ might be partly because much of the energy of /l/ is in the lower frequencies (though by no means always mainly in F_1). This explanation seems unlikely to entirely account for the findings, however, since S1's clear onset /l/ has rather high frequencies, and her F_1 in onset /l/ was not affected by the voicing of the coda (Figures 1 and 2). Appendix 2 discusses the COG further with respect to fricatives.

3.6 Spectral shape: Formant transitions

Formant transitions in four words with vowel /ɛ/ in Speaker 1's data were analysed for other purposes. We report the results here for their relevance to previous work in the literature. When the words were spoken in isolation (but not when they were in the carrier phrase), F_1 transitions at the beginning of the vowel rose less steeply and lasted 10-20 ms longer when the coda was voiced. Likewise, F_2 transitions appeared slower when the coda was voiced. Carter (1999; pers.comm.) similarly observed that longer onset /l/ seems to have slower transitions into the vowel, at least when the initial /l/ is clear. These coda-dependent differences, which may be reflected in the spectral patterns observed at vowel onset (V-on in Figures 1 and 2), mirror the pattern at the other end of the syllable: Chen (1970) and Fujimura & Miller (1979) have shown that F_1 transitions at a VC boundary are slower when the C is voiced, at least for low vowels. These small and relatively simple patterns in the acoustics may mask quite complex production effects: a number of experiments have shown that onset transitions in vowels preceding voiced vs. voiceless stops show complex effects in which tongue, jaw, and probably lip movements seem to be quasi-independent so that their combined effects produce rather small acoustic differences overall (Wolf, 1978; Fitch, 1981; Summers, 1987; Löfquist & Gracco, 1994).

3.7 Fundamental frequency

Although Hawkins and Nguyen's (2000, 2001) perceptual experiments show that listeners labelled truncated synthetic syllables as *led* when the f_0 in the /l/ was low and as *let* when f_0 was higher, coda voicing in the natural speech of the present study had no systematic effect on f_0 in the /l/. Table VI shows that of the eight f_0 differences measured at the midpoint and the end of /l/ (four speakers x two locations), six were of less than 1 Hz, and the other two were less than 3 Hz (+2.9 Hz and –1.5 Hz). At V-on and V-mid, the range of coda-dependent f_0 differences likewise varied between only –2 Hz and +3 Hz, apparently randomly. The lower f_0 at V-end (30 ms before the end of the vowel) when the coda was voiceless, which was observed for S1-S3, probably reflects glottalisation before a voiceless coda; in other words, f_0 cannot be properly measured at this location when the coda is voiceless.

TABLE VI ABOUT HERE

TABLE VII ABOUT HERE

Although some of these differences were statistically significant, as indicated by asterisks in Table VI (and see Table VII for the F ratios), we believe that all except those at the end of the vowel should be disregarded in the present study. Our reasons are twofold, one methodological and one perceptual. First, the speakers in this experiment were required to use as constant a pitch range as possible because some of the utterances were to be cross-spliced for the lexical decision task reported by Hawkins and Nguyen (in press). Although this requirement might have reduced the size of a natural coda-dependent difference in f0, it seems very unlikely to have done so. To take an analogy, we are not aware that the requirement to speak on a monotone significantly reduces vowel-dependent differences in intrinsic f0. It seems much more likely that the requirement to use the same pitch contour produced artificially-small variances in f0, with consequent statistically-significant differences in a few of the mean values. Second, it seems unlikely that the observed f0 differences would be perceptually salient in these stimuli. Relative pitch in the vicinity of the obstruent closure is probably a perceptual cue to coda voicing (cf. Kingston & Diehl 1995), and such cues may be generalised to other parts of the syllable in the absence of other strong cues (Hawkins & Nguyen 2000, 2001) but differences as small as 0.5–3 Hz are close to the jnd for pitch of sinusoidal tones at these frequencies, and thus seem unlikely to be detectible in natural speech with a normal range in the intonation contour, as used in this experiment. The fact that the significant differences are distributed unsystematically across speakers, speaking conditions, and syllable locations makes them even less likely to function as usable perceptual cues (cf. Hawkins & Nguyen 2001). However, this issue might bear more study in spontaneous speech, for it is striking that listeners consistently associate low f0 with coda voicing (Hawkins & Nguyen 2000, 2001).

3.8 *Correlations* between *duration of onset /l/ and F_2 frequency at its midpoint*

For each speaker, Pearson correlations were calculated between the duration of /l/ and (1) the frequency of F_2 at /l/ midpoint and (2) the COG at /l/ midpoint in each monosyllable. All correlations were close to zero.

4 Discussion

This study confirms that when the syllable coda is voiced, a syllable-onset /l/ as well as the nucleus is longer than when the coda is voiceless. Data presented in Appendix 2, which show that S1's onset voiceless fricatives are also longer before a voiced coda, raise the possibility that this coda-dependent lengthening of onsets may not be restricted to onsets that are sonorant. At least when the onset is an /l/, there are typically spectral as well as durational differences in the /l/ and/or the following vowel. All four speakers produced diverging patterns of differences in F_1 and F_2 frequencies when the syllable coda was voiced relative to when it was voiceless. F_1 frequencies evolved towards relatively lower F_1 before voiced codas as they progressed from the onset /l/ through to the midpoint of the vowel, while F_2 frequencies evolved towards increasingly higher frequencies towards V-mid. This pattern supports the findings of Wolf (1978) and Summers (1988), particularly those of Wolf, and extends them to include the syllable onset as well. Some particular vowel qualities will produce different patterns. For example, our data confirm Kwong & Stevens' (1999) prediction that rounded vowels may produce the opposite changes in F_1 frequency. Note, however, that our observed differences pattern with those of Wolf (1978) and Summers (1988) rather than those of Kwong & Stevens; the reasons are unclear, but might include differences due to monosyllabic vs. disyllabic word structure, and the fact that Kwong & Stevens' data involved (phonetic) taps rather than (phonetic) stops.

For most speakers, the significant spectral differences are not evenly distributed, but localized to a particular region of the sonorant+vowel sequence. Differences in F_1 were mainly confined to the vowel for S1, S2, and S3, except for S2's back-rounded vowel group. S2 and S3 confine their largest and most consistent F_2 differences to the onset /l/, whereas S1 makes it in the vowel (Figure 2). (These observations neglect V-end in the COG, centred 30 ms before the end of the vowel (Figure 4), which is affected by well-known differences at vowel offset (abrupt vs. gradual cessation of periodicity), dependent on coda voicing.) These trends are broadly consistent with Slater & Coleman's (1996) observation that differences associated with stop voicing affect the preceding sonorant+vowel sequence unevenly, and may be confined to a specific temporal region of this sequence.

The particular F_2 pattern produced appears to depend on both phonological and phonetic properties of the speaker's accent. The critical phonological contrast appears to be whether onset /l/ is darker or lighter than onset /r/. The more common phonological contrast for accents of British and American English is for onset /l/ to be lighter than onset /r/. In these accents, represented by S1, S2, and S3, the data suggest that onset /l/s will be relatively long and dark (or at least with a lower center of gravity) when the coda of the same syllable is voiced, and relatively short and light when the coda is voiceless. Less common are accents in which onset /l/ sounds darker than onset /r/. S4, the one subject in the current study whose accent had this pattern, produced onset /l/s that were longer but not darker before voiced codas; in fact, S4's F_2 was nonsignificantly higher before voiced codas in the /l/, and significantly higher at vowel onset.

The spectral differences dependent on the phonetic quality of /l/ suggest that accents with clear initial /l/ may 'darken' the syllable by lowering the spectral center of gravity in some way without significantly backing the tongue. Although in this study S1 was the only representative of an accent with such a 'clear' phonetic quality to onset /l/ in this study, this clear onset /l/ is the form used in standard British English, as well as in a number of other regional accents, and so this issue merits closer investigation with more speakers.

4.1 Coda voicing as an attribute of the whole syllable: 'sombre~bright'

The main contribution of this study to the literature on acoustic properties associated with the voicing distinction of syllable-coda obstruents is summarised in Figure 5, which shows spectrograms of tokens of *led* and *let*, spoken by S1, together with spectra (14-pole lpc, 50-ms $\cos^4$ window) centred at the midpoint of the /l/. The frequency of F_2 in /l/ is relatively higher than that found for onset /l/ in the speech of S2, S3, S4, and speakers of American English, and is typical for the phonetically clear onset /l/ of S1's accent. Letters a, b, and c refer respectively to vowel duration, the F_1 vowel offset transition, and the F_1 vowel onset transition. Coda-dependent differences associated with these parts of the syllable have been observed by others as well as ourselves, as noted above. The longer vowel duration before voiced obstruent codas, along with the long F_1-offset transition culminating in lower F_1 frequency at vowel offset, have been consistently found in the literature. In comparison, coda-dependent variations in F_1 onset transition are not well understood: they may be less consistent, or perhaps dependent on other properties of the syllable such as the particular segments in the onset, the vowel quality, and whether the word is spoken in citation form, as noted in Section 3.6. Clearly visible in Figure 5, but unlabelled, are distinguishing attributes of the stop segment itself whose perceptual importance is well established: for voiced obstruent codas, the short closure duration, and the low amplitude and low peak frequency range of the burst (or, more generally, of the aperiodicity).

FIGURE 5 ABOUT HERE

The principal original contribution of this study is in the coda-dependent differences in the syllable onset: when the coda is voiced, the duration of onset /l/s, and probably of fricatives (see Appendix 2), is longer, and F_2 and the COG of onset /l/ are usually lower. In Figure 5, /l/ is 148 ms in *led*, compared with only 125 ms in *let*; F_2 frequency is 1750 Hz in *led* and 1875 Hz in *let*. The corresponding COG measures, which are not shown on the Figure, are 831 Hz and 889 Hz. The utterances shown in Figure 5 have the same F_1 frequency at the midpoint of the /l/, which is representative for S1 and S2 with other non-round vowels (see Figure 2), and for S3 for all vowels (Figure 1). There may however be a tendency for F_1 to be lower in onset /l/ before voiced codas, and this would be expected to be reliable for S4's accent, and for back rounded vowels in at least some other accents. In the vowel, the slowly rising F_2 of *led* is clearly visible (as noted in Section 3.6). In comparison, F_2 of *let* starts higher and stays relatively steady. This coda-dependent difference between the two F_2 trajectories may be a primary cause of the gradual change in the balance of F_2 between V-on and V-mid (Figures 1 and 2, Section 3.4; see also Footnote 2).

In summary, when a syllable has a voiceless coda, the onset and nucleus are shorter, and often start phonetically lighter, or brighter; the coda is also brighter. Impressionistic observation, supported by our own measurements, suggest that this is the pattern for most accents of (British?) English; other patterns are found in some accents, as evidenced by S4 in this study. For simplicity, most of the rest of this discussion focuses on the most common pattern.

We use the terms *bright* and *sombre* to describe the endpoints of a perceptual dimension that reflects the complex sets of acoustic properties under discussion. We could have chosen to use [±voice], or *clear/light* vs. *dark*, but each of these already has other meanings which we wish to distinguish. There are many precedents for a descriptor involving brightness (cf. Butcher 1974), and Fischer-Jørgensen (1985) has argued for a single perceptual dimension of brightness/darkness to describe complex acoustic correlates of vowel quality. Our acoustically-complex perceptual dimension can apply to whole syllables (to consonants as well as vowels) and we have substituted *sombre* for *dark*, in order to avoid confusion with the narrower impressionistic phonetic term *dark*, which is often applied to /l/ and typically equated with tongue dorsum backing and/or velarisation (the two not being easily distinguished auditorily). Thus we shall continue to use *dark* and *light/clear* with their normal phonetic meanings, and *voiced/voiceless* to mean phonated or not (phonetically) and phonemes such as /b d g/ vs. /p t k/ phonologically. As we shall argue, fortis articulations will tend to produce perceptually bright sound, and lenis articulations will tend to produce perceptually sombre sound. (Correspondences are much more tenuous between dark and lenis, and between light and fortis, and we make no claims about any such correspondences.) Finally, a perceptual dimension that can encompass a number of distinct articulatory and acoustic correlates is appealing, because it makes no untested claims about articulation.

Our findings suggest that voicing in the coda defines the nature of the whole syllable: for roughly the same overall duration, the spectrogram of *led* in Figure 5 is characterized largely by sonorance, and thus in some sense sombreness, whereas that of *let* is characterized by a high proportion of silence or aperiodicity—in some sense brightness. These rather obvious differences between the two types of syllable are accompanied by small but systematic spectral differences. Specifically, the spectral pattern for *led* begins and ends with a relatively sombre quality, passing through a relatively brighter stage in the middle of the vowel and into its second half; this mid-syllable brightness might enhance the sombre quality of the syllable edges. In contrast, the periodic parts of *let* change less in brightness between the middle of the syllable and its edges, and these voiceless-coda syllables thus begin and end relatively brighter than those with voiced codas[2]. Taken together with the very different durations of the phonated portions of the two syllables, then, the overall impression is that voiceless-coda syllables are probably experienced as relatively brighter than voiced-coda syllables in the accents represented by S1, S2, and S3.

If this analysis is accepted, then [±voice] in the coda can be described as a property of the whole syllable, though the bright~sombre dimension is perhaps a better description. When the coda is voiceless, this property manifests as relatively bright: small proportions of relatively high-frequency periodic energy, and a high proportion of silence or aperiodic energy which may also be relatively light. When the coda is voiced, this property manifests as relatively sombre: a high proportion of relatively low-frequency periodic energy at the syllable edges (contrasting with the middle of the nucleus) and relatively small amounts of silence and aperiodic energy.

This description is inconsistent with Weismer's (1979) finding that syllable-onset VOT is slightly longer before voiced than voiceless codas. However, Weismer's results may be explained by aerodynamic advantages due to lower F_1 before voiced codas, as found by others. Noting the plausibility of an aerodynamic explanation, Weismer concluded that his observed differences were unlikely to be perceptually salient (they averaged 3-5 ms, rising to 12 ms when the vowel was tense). However, in view of our own perceptual findings summarised above, it seems worth speculating whether such small but systematic differences might have a cumulative perceptual effect. It would also be worth measuring the duration of the stop closure preceding the VOT, in view of our own findings for fricatives, and since perceptual responses seem likely to be influenced by the relative durations of the closure and VOT. Pending further information, we conclude that Weismer's results do not invalidate an interpretation in terms of a sombre~bright perceptual dimension.

Finally, notice that S4's voiced-coda syllables are consistently bright (apparently relatively fronted) throughout, relative to his voiceless-coda syllables. This contrasts with the sombre-bright-sombre pattern of S1, S2, and S3. This comparison is of course within each speaker's own utterances: syllables with voiced codas relative to those with voiceless ones. In absolute phonetic quality, all S4's and S3's syllables are very dark; S2's are rather lighter but still somewhat dark; S1's are all light, or clear. This contrast, with S4 using brighter syllables when the coda is voiced, and the other three using the sombre-bright-sombre pattern through the voiced-coda syllable, thus neatly reflects the phonological difference between S1, S2, and S3 on the one hand (relatively darker onset /l/ than onset /r/) and S4 on the other (relatively darker onset /r/ than onset /l/). This potential contribution of a complex auditory percept to lexical distinctiveness (in dialect-specific ways) seems a worthwhile area for further investigation. For example, it could be worth examining the relative sombreness of voiced and voiceless coda stop bursts in these or similar accents. Whereas accents represented by S1, S2, and S3 would be expected to have /t/ brighter than /d/, one might expect accents represented by S4 to have /d/ bursts brighter than /t/ bursts.

4.2 Articulatory correlates

This section discusses the coda-dependent spectral properties in the onset from two points of view. First, we bring together and extend our earlier arguments about what types of articulatory manoeuvre could produce the observed spectral differences; we conclude that the best explanation for the /l/ seems to be tongue dorsum backing (TDB), which may be velarisation in many cases. Second, we consider whether these onset properties are independent of the physical adjustments needed to produce voiced and voiceless obstruents, or whether they result from anticipatory coarticulation of an essential attribute of the coda. If TDB/velarisation is an essential or an enhancing attribute of coda voicing, then these data could be interpreted as a subtle example of anticipatory spread of coarticulation, like nasalization and lip rounding. As such, they would be interesting but possibly not worth much further study. On the other hand, if TDB/velarisation is not integral to coda obstruent voicing, then there would be far-reaching theoretical implications, as developed below.

We have noted that the spectral patterns capture both the phonological and phonetic differences between speakers, with S1, S2, and S3 forming one phonological group, distinct from S4, while S1 is distinct from S2 and S3 in maintaining an equally clear /l/ before both voiced and voiceless codas. S1 and S4 sound very different in the absolute phonetic qualities of the /l/s: S4's onset /l/ is very dark, while S1's is clear.

We have suggested that the best explanation for the majority of these accents seems to be that a voiced coda engenders tongue dorsum backing in the onset /l/, and possibly fronting and/or jaw/tongue lowering in the vowel. Absence of reliable changes in F_1 during the /l/ suggests that there is rarely jaw/tongue lowering in onset /l/. Could the F_2-lowering observed for S2 and S3 in the onset /l/ be velarisation, i.e. tongue dorsum raising as well as backing? The coda-dependent spectral differences observed in onset /l/ are consistent with velarisation, in that velarisation is thought to affect the frequency of F_2 and not F_1 (Stevens, 1998: 543-554). Perhaps the small and unreliable changes in F_1 could also arise from velarisation: Stevens models F_1 of /l/ as a Helmholtz resonance, so conceivably a more velarised /l/ could be accompanied by small changes in F_1 frequency, in that backing and raising the tongue dorsum will change the volume of the oral cavity contributing to the acoustic compliance component of the Helmholtz resonance, and possibly also of the acoustic mass, via small changes in the length and cross-sectional area near the alveolar closure. However, these speculative comments need corroborative movement data. Without them, it seems clear that tongue backing is involved, but it remains an open question whether the articulation involves simply backing the dorsum, or raising it, or more complex manoeuvres including, perhaps, changes in larynx height. Much might depend on the 'unmarked' degree of velarisation in the particular accent, and the quality of the following vowel. Thus, it could be worthwhile to see if high vowels engender velarisation while low vowels do not, in that, when the tongue tip is anchored on the alveolar ridge as for an onset /l/, raising the dorsum might be a simple way to produce the desired acoustic-perceptual effect without unduly distorting the quality of the following vowel; conversely, TDB without raising might be preferred in onset /l/ before low vowels.

The above argument assumes that TDB/velarisation are 'optional additions' to an /l/ in certain accents. However, Sproat & Fujimura (1993) argued from articulatory evidence that the main cause of the correlation between darkness and duration for post-vocalic /l/ is coarticulatory undershoot: when /l/ is short, the tongue has less time to take up a more extreme position. In our view, this explanation does not account convincingly for our data, for there seems no reason to assume coarticulatory undershoot in *onset* /l/s that are clearly spoken in citation form, particularly when both the /l/ and the vowel are longer. Moreover, many of the onset /l/s produced by S2, S3, and S4 were both very short and very dark, with S4's /l/s (nonsignificantly) darker when they were shorter (i.e. before voiceless codas); S1 had the longest onset /l/ durations, and a phonetically clear /l/. Huffman (1997) likewise concluded that the relationship between duration and backness of onset/intervocalic /l/ articulations in American English is variable, and that the factors affecting it are considerably more complex than Sproat & Fujimura (1993) suggested. On the other hand, we believe our data are compatible with the main claims of Sproat & Fujimura's model, and we strongly endorse their position that /l/ can be produced with a continuum of darkness, as can be demonstrated by moving the tongue body around while keeping the tip on the alveolar ridge and the sides in the right configuration to produce a lateral. From these and other observations, and perhaps in particular S4's data, we conclude that there may be rather a lot of freedom in the relative timing of apical and dorsal gestures of /l/. Carter (2002) also argues for this articulatory freedom, and suggests that it is exploited by non-rhotic accents of English, with different accents arbitrarily using one pattern for /l/, and the opposite one for /r/, thus setting up a systematic opposition between clear and dark liquids (bright and sombre in our terminology) in the phonological domain.

We turn now to the second issue, which is to consider the evidence for tongue dorsum backing (or similar articulatory manoeuvres) being somehow integral to coda obstruent voicing. Superficially, one could conclude that tongue dorsum backing cannot be integral to coda voicing because the accent represented by S4 behaves differently from those of the other three speakers. However, that argument is unsatisfactory, because the accents of Speakers 1, 2, and 3 might capitalise on a natural affinity between coda voicing and tongue dorsum backing which S4's accent has overcome for independent reasons.

One possibility is that coda voicing is facilitated by changes in the volume of the oral cavity. Increasing the volume of the oral cavity could facilitate physical voicing in the coda by maintaining a greater transglottal pressure drop, and Westbury (1983) observed oral-cavity enlargement during the closure periods of phonologically voiced stops that were also phonetically voiced (i.e. phonated). However, Westbury (1983) and Stevens (1998, see also Kwong and Stevens 1999) argue that such enlargement must take place during the obstruent constriction; anticipatory enlargement might be more likely to facilitate voicelessness by stretching the vocal tract tissue (which might include stiffening the vocal folds) and making it difficult to actively or passively expand the oral cavity further during an obstruent articulation. It seems unlikely that we will find a satisfactory answer to these issues while we lack relevant articulatory data. In the absence of directly relevant data, perhaps the most convincing evidence that the observed differences in onset /l/ do not result from essential attributes of coda voicing is that the pattern for syllables with voiced codas changes progressively from the /l/ into the vowel midpoint; opposite patterns in the onset and the vowel seem incompatible with an interpretation in terms of anticipatory coarticulation throughout the syllable.

Another possibility is that vocal-tract tension might be controlled to produce fortis voiceless codas and lenis voiced codas. In standard phonetic theory, fortis-lenis is generally noted as an enhancing rather than an essential dimension of obstruent voicing in English, so its spread is not best ascribed to anticipatory coarticulation of essential attributes of stop articulation. However, one could adopt the view that the voicing distinction for coda stops is more correctly analysed as a fortis-lenis distinction. The acoustic and articulatory correlates of the fortis-lenis distinction are complex and poorly understood, but it seems reasonable to assume that their effects can be manifest over several segments, and that they contribute to the complex bright~sombre dimension, with fortis articulations producing relatively bright syllables and lenis ones producing sombre syllables.

If we adopt this reasoning, our data suggest that when a coda is voiced (lenis), the entire syllable, not just its rhyme, is characterized by lenis properties. These adjustments may have developed to facilitate coda voicing, but whatever their historical antecedents, our observed patterns do not point to anticipatory coarticulation of essential coda properties, but rather to a complex, syllable-wide acoustic-phonetic realisation of a simple phonological distinction.

Why does it matter whether spectral differences in the /l/ reflect essential attributes of coda voicing, or merely accompany the distinction? Whereas the durational pattern simply extends the well-established distinction in vowel length due to coda voicing into another sonorant part of the syllable, the association of sombreness with coda voicing potentially adds a different dimension. The properties we have identified are in non-adjacent phonetic segments. If the onset properties are independent of the coda properties, then this case is quite different from better-known instances of nonadjacent influences, such as vowel-to-vowel coarticulation, spread of lip-rounding, or nasalization, in which essential attributes of the influencing segment spread. The next section summarises some of the theoretical implications of this conclusion for models of speech perception.

4.3 Implications for perception

These data show that the voicing of a syllable's coda is reliably reflected in its onset, and other experiments, cited above, show that those acoustic properties help listeners to determine the voicing of the coda. The onset /l/ provides only weak perceptual cues to coda voicing, but even weak cues can be useful, for example by enhancing an otherwise impoverished stimulus in adverse listening conditions (*cf.* Whalen, 1989; Pisoni & Lively, 1995). They may be especially salient when they are distributed over longer time domains than the stronger primary cues, in that they could contribute to the signal's overall perceptual coherence (Hawkins, 1995; Pisoni, 1991; Remez, Rubin, Berns, Pardo & Lang, 1994; Ogden, Hawkins, House, Huckvale, Local, Carter, Dankovicová, & Heid, 2000; Remez, 2001) as well as allowing the listener longer to process particular linguistic properties (*cf.* Warren & Marslen-Wilson, 1987; Marslen-Wilson & Warren, 1994).

These points confirm observations from a number of experiments conducted over at least the last two decades, as noted earlier (see also Alfonso & Baer, 1982; Fowler & Smith, 1986) but two aspects of the present data go further than previous work: the acoustic-perceptual correlates of coda voicing appear in a phonetic segment that is *not adjacent* to the conditioning segment and may not appear in the same way in the intervening segment (the syllable nucleus); and these correlates are not readily ascribable to anticipatory coarticulation of an essential property of the coda in current pronunciations, and may indeed be realisations of part of the phonological system of liquids, at least for nonrhotic accents of English (Carter 2002).

Although these data can be accounted for by models of speech perception that assume a basic phoneme-sized unit of perception, they seem to us to lend stronger support to nonsegmental models of speech perception and word recognition. In nonsegmental models, the detailed fine structure of the incoming speech signal is mapped directly onto lexical and grammatical representations or indeed directly onto a representation of meaning, and segmental or phonemic representations may not be involved at all in the process of understanding an utterance. Such models emphasize the potential perceptual role of fine phonetic detail that may be localised to a particular part of the signal or distributed over long domains. This view is discussed in Nguyen and Hawkins (1999), Hawkins and Smith (2001), Hawkins & Nguyen (in press) and Hawkins (submitted), and is compatible with work by Streeter & Nigro (1979), Suomi (1993), Marslen-Wilson & Warren (1994), Jusczyk (1997), and Whalen, Best & Irwin (1997), amongst others, as well as with the motivating premises of some recent computational models of word recognition (e.g. Guenther & Gjaja, 1996; Gaskell & Marslen-Wilson, 1997; Plaut & Kello, 1999; Protopapas, 1999). Together with other work on very long-domain liquid resonance effects (Hawkins & Slater, 1994; Tunley, 1999; West, 1999, 2000; Heid & Hawkins, 2000), these data suggest that speech perception must operate with (at least) two time windows: a short one to analyse short-time events, which are often highly informative (*cf.* Stevens, 1983, 1998) and a much longer one, possibly extending to several syllables, which may track the signal's coherence and the weak but nevertheless salient long-domain segmental information demonstrated here. How information from these long and short temporal windows is combined is a challenging problem for the future (*cf.* Grossberg, Boardman & Cohen, 1997; Grossberg & Myers, 2000).

In summary, if information about coda voicing is available to listeners in the syllabic onset, then coda voicing could reasonably be modelled as a property of the entire syllable. This view is consistent both with the position that words can be recognised from relatively weak auditory information spread across more than one acoustic-phonetic segment, as long as it is consistent, and with the stronger claim that speech is analysed and matched in a nonsegmental way rather directly to meaning, or at least to lexical items, which might themselves be represented nonsegmentally.

Acknowledgements

Funded in part by grants from the Swiss National Science Foundation (83BC-056471) and the EPSRC (GR/N19595). We thank our speakers, Richard Steadman-Jones, Gavin Burnage, and Jaime Craig; Sebastian Heid for help with labelling; John Coleman and John Local for valuable discussion; and Bryan Gick, Doug Whalen, and two anonymous reviewers for helpful suggestions for improving the manuscript. Parts of this paper were presented at ICSLP-98, Sydney, Australia (Nguyen & Hawkins, 1998).

Appendix 1: Experimental syllables used in the perception experiment

col. 1: whether the word ends in a voiced coda (VO) or a voiceless one (VL)
col. 2: vowel category (back rounded; high front; low-back or central; low front)
col. 3: (C)lVC words
col. 4: frequency of occurrence of the (C)lVC words
col. 5: number of high-frequency monosyllabic competitors for the (C)lVC word
col. 6: (C)lVC nonword
col. 7: FVC words
col. 8: frequency of occurrence of the FVC words
col. 9: number of high-frequency monosyllabic competitors for the FVC word
col. 10: FVC nonword

(1) coda	(2) vowel categ.	(3) (C)lVC word	(4) freq.	(5) compet-itors	(6) (C)lVC nonword	(7) FVC words	(8) freq.	(9) compet-itors	(10) FVC nonword
VO	br	blob	1	0	blop	sod	1	0	sot
VO	br	glued	12	0	gloot	chewe	3	0	chewt
						d		1	
VO	br	load	44	1	loat	sowed	2	0	sote
VO	br	slog	1	0	slock	fog	12	0	fock
VL	br	bloke	4	0	blogue	choke	5	0	chogue
VL	br	blot	3	0	blod	shock	53	1	shog
VL	br	cloak	10	2	cloague	soak	0	0	soag
VL	br	flop	2	0	flob	chop	4	1	chob
VL	br	luke	7	0	loog	soup	10	0	sube
VL	br	plop	2	0	plob	shop	84	0	shob
VL	br	slope	18	0	slobe	soap	4	1	sobe
VO	hf	blade	18	0	blate	shade	30	2	shate
VO	hf	glib	0	0	glip	fib	0	1	fip
VO	hf	glide	1	0	glite	chide	1	1	chite
VL	hf	bleep	0	0	bleeb	sheep	27	0	sheeb
VL	hf	blip	0	0	blib	ship	57	1	shib
VL	hf	clique	2	0	cleague	cheek	22	1	cheeg
VL	hf	flake	0	0	flague	fake	5	2	fague
VL	hf	slake	0	0	slague	sake	44	1	sague
VL	hf	sleet	4	0	sleed	cheat	8	2	chead
VL	hf	slick	1	0	slig	sick	49	1	sig
VO	lb/c	blurb	1	0	blurp	Serb	0	0	serp
VO	lb/c	flood	21	0	flut	thud	6	0	thut
VO	lb/c	glug	0	0	gluck	chug	1	0	chup
VO	lb/c	lard	0	2	lart	shard	0	0	shart
VL	lb/c	flirt	1	0	flird	shirt	29	1	shirp
VL	lb/c	flute	3	0	flewed	suit	55	0	soode
VL	lb/c	glut	0	0	glud	shut	39	0	shud
VL	lb/c	lark	6	2	larg	shark	0	0	sharg
VL	lb/c	lurk	1	1	lurg	shirk	1	1	shirg
VL	lb/c	slurp	0	0	slurb	chirp	0	2	chirb
VL	lb/c	slut	0	0	slud	suck	1	1	sug
VO	lf	blab	0	1	blap	fab	0	0	fap
VO	lf	blared	0	0	blairt	shared	42	0	shairt
VO	lf	fled	3	0	flet	shed	21	3	shet
VL	lf	blight	2	0	blide	fight	87	0	fide
VL	lf	clack	3	0	clag	chat	10	1	chab
VL	lf	clap	1	0	clab	sap	1	0	sab
VL	lf	fleck	0	0	fleg	check	58		cheg

Appendix 2
Acoustic properties of onset fricatives and affricates dependent on coda voicing

The speech recorded for the lexical decision task (Hawkins & Nguyen, in press) included 39 FVC syllable, spoken by S1, where F is a fricative or affricate and the two members of each pair differ in the voicing of the final stop. The pairs were matched with the (C)lVCs in terms of vowel quality, word frequency and number of competitors, as shown in Appendix 1. Their onset durations and spectral shape were measured as a first attempt to see if the coda-dependent differences we observed in onset /l/ are likely to be restricted to onset sonorants, or might apply to a wider set of sounds. The durations of these obstruents were measured from the onset of aperiodic energy to its offset, excluding any regions of mixed aperiodic and periodic excitation, and the spectral COG was measured as in the main experiment, centred at the midpoint of the aperiodic noise.

Like onset /l/, these onset fricative potions were slightly longer before voiced than before voiceless codas: 368 ms vs. 351 ms ($F(1,35) = 14.37$, $p = 0.001$; $F(1,34) = 14.61$, $p = 0.001$, for vowel and syllable-onset ANOVAs respectively). There were no significant interactions. The COG showed expected vowel-dependent differences, but no significant coda-dependent differences.

Although these data are sparse, and only from one speaker, they are nevertheless noteworthy in that they suggest that lengthening before a voiced coda may not always be restricted to sonorant parts of syllables. If this finding proves reliable, it will have important consequences for how we view the articulatory-acoustic structure of words. The lack of difference found for the spectral measures may accurately reflect that there are no differences in fricatives commensurate with those found for onset /l/. On the other hand, S1's spectral differences in onset /l/ were also small: other speakers may thus have larger differences in onset fricatives. Moreover, although the COG has been successfully used to differentiate coarticulatory distinctions amongst fricatives (*cf.* Jassem, 1979; Forrest, Weismer, Milenkovic & Dougall, 1988; Hoole, Ziegler, Hartmann & Hardcastle, 1989; Hoole, Nguyen & Hardcastle, 1993; Nittrouer, Studdert-Kennedy & McGowan, 1989), our observations suggest that, just as with the vowels in syllables with onset /l/, it may need modification if it is to be sensitive enough to reflect the subtle differences under study here. When the COG is calculated over a wide amplitude range, small differences in the shape of the low-amplitude skirts of spectral peaks can influence the value of the COG in ways that probably exaggerate their auditory influence; furthermore, random fluctuations in noise excitation may produce spurious spectral peaks that contribute noise to the COG measurements, despite the methods used here to avoid such undue influence. Either of these sources of variation in the data is likely to be compounded by the already wide variation in spectral detail due to the range of vowels used in this study. A number of modifications to the general method might thus be worthwhile, including (1) restricting the amplitude and possibly the frequency range (*cf.* Beddor & Hawkins, 1990), and (2) comparing COGs in different frequency ranges. These issues for obstruent onsets, together with measures of formant frequency in the vowels, would be worth pursuing on a purpose-designed data set from more speakers.

References

Alfonso, P. & Baer, T. (1982) Dynamics of vowel articulation. *Language & Speech* **25**, 151-173.

Beddor, P.S. & Hawkins, S. (1990) The influence of spectral prominence on perceived vowel quality. *Journal of the Acoustical Society of America* **87**, 2684-2704.

Butcher, A. (1974) Brightness, darkness, and the dimensionality of vowel perception. *Journal of Phonetics* **2**, 153-160.

Carter, P. (1999) Abstractness in phonology and extrinsic phonetic interpretation: The case of liquids in English. In *Proceedings of the XIVth international congress of phonetic sciences* (J.J. Ohala, Y. Hasegawa, M. Ohala, D. Granville, *&* A.C. Bailey, editors), University of California, Berkeley, CA., **1**, 105-108.

Carter, P. (2002) Structured variation in British English liquids: The role of resonance. PhD dissertation, University of York.

Chen, M. (1970) Vowel length variation as a function of the voicing of the consonant environment. *Phonetica* **22**, 129-159.

Denes, P. (1955) Effect of duration on the perception of voicing. *Journal of the Acoustical Society of America* **27**, 761-764.

Fischer, R.M. & Ohde, R. (1990) Spectral and duration properties of front vowels as cues to final stop-consonant voicing. *Journal of the Acoustical Society of America* **88**, 1250-1259.

Fischer-Jørgensen, E. (1985) Some basic vowel features, their articulatory correlates, and their explanatory power in phonology. In *Phonetic linguistics* (V. A. Fromkin, editor), pp.79-99. Orlando, FL: Academic Press.

Fitch, A. (1981) Distinguishing temporal information for speaking rate from temporal information for intervocalic stop consonant voicing. *Haskins Laboratories Status Report on Speech Research* **65**, 1-32.

Fowler, C.A. & Smith, M.R. (1986) Speech perception as 'vector analysis': an approach to the problem of invariance and segmentation. In *Invariance and variability in speech processes* (J.S. Perkell & D.H. Klatt, editors), pp. 123-136. NJ: Lawrence Erlbaum.

Forrest, K., Weismer, G., Milenkovic, P. & Dougall, R.N. (1988) Statistical analysis of word-initial voiceless obstruents: preliminary data. *Journal of the Acoustical Society of America* **84**, 115-123.

Fujimura, O. & Miller, J.L. (1979) Mandible height and syllable-final tenseness. *Phonetica* **36**, 263-272.

Gaskell. M.G. & Marslen-Wilson, W.D. (1997) Integrating form and meaning: A distributed model of speech perception. *Language & Cognitive Processes* **12**, 613-656.

Grossberg, S. & Myers, C.W. (2000) The resonant dynamics of speech perception: Interword integration and duration-dependent backward effects. *Psychological Review* **4**, 735-767.

Grossberg, S., Boardman, I. & Cohen, M. (1997) Neural dynamics of variable-rate speech categorization. *Journal of Experimental Psychology: Human Perception & Performance* **23**, 481-503.

Guenther, F.H. & Gjaja, M.N. (1996) The perceptual magnet effect as an emergent property of neural map formation. *Journal of the Acoustical Society of America* **100**, 1111-1121.

Hawkins, S. (1995) Arguments for a nonsegmental view of speech perception. In *Proceedings of the XIIIth international congress of phonetic sciences* (K. Elenius & P. Branderud, editors), **3**, 18-25. Stockholm: KTH and Stockholm University.

Hawkins, S. (submitted) Roles and representations of systematic fine phonetic detail in understanding speech. *Journal of Phonetics.*

Hawkins, S. & Nguyen, N. (2000) Predicting syllable-coda voicing from the acoustic properties of syllable onsets. In *SWAP: The ISCA conference on spoken word access processes* (A. Cutler, J. McQueen & R. Zondervan, editors), pp.167-170. Nijmegen: Max-Planck-Gesellschaft zur Förderung der Wissenschaften.

Hawkins,S, and Nguyen, N. (2001) Perception of coda voicing from properties of the onset and nucleus of *led* and *let*. In P. Dalsgaard, B. Lindberg, and H. Benner (eds.) *Proceedings of the 7th International Conference on Speech Communication and Technology (Eurospeech 2001 Scandinavia)* 1, 407-410.

Hawkins, S. & Nguyen, N. (in press) Effects on word recognition of syllable-onset cues to syllable-coda voicing. In *Papers in laboratory phonology VI* (J.K. Local, R.A. Ogden & R.A.M. Temple, editors). Cambridge University Press.

Hawkins, S. & Slater, A. (1994) Spread of CV and V-to-V coarticulation in British English: Implications for the intelligibility of synthetic speech. *Proceedings of the 3rd international conference on spoken language processing* **1**, 57-60.

Hawkins, S. & Smith, R. (2001) Polysp: A polysystemic, phonetically-rich approach to speech understanding. *Italian Journal of Linguistics—Rivista di Linguistica* 13, 99-188.

Heid, S. & Hawkins, S. (2000) An acoustical study of long-domain /r/ and /l/ coarticulation. *Proceedings of the 5th Seminar on Speech Production: Models and Data.* (ISCA), pp.77-80. Kloster Seeon, Bavaria, Germany.

Hogan, J.T. & Rozsypal, A.J. (1980) Evaluation of vowel duration as a cue for the voicing distinction in the following word-final consonant. *Journal of the Acoustical Society of America* **67**, 1764-1771.

Hoole, P., Nguyen, N. & Hardcastle, W.J. (1993) A comparative investigation of coarticulation in fricatives: electropalatographic, electromagnetic and acoustic data. *Language & Speech* **36**, 235-260.

Hoole, P., Ziegler, W., Hartmann, E. & Hardcastle, W.J. (1989) Parallel electropalatographic and acoustic measures of fricatives. *Clinical Linguistics & Phonetics* **3**, 59-69.

House, A.S. & Fairbanks, G. (1953) The influence of consonant environment upon the secondary acoustical characteristics of vowels. *Journal of the Acoustical Society of America* **25**, 105-113.

Huffman, M. (1997) Phonetic variation in intervocalic onset /l/'s in English. *Journal of Phonetics* **25**, 115-141.

Jassem, W. (1979) Classification of fricative spectra using statistical discriminant functions. In *Frontiers of speech communication research* (B. Lindblom & S. Öhman, editors), pp.77-91. London: Academic Press.

Jones, D. (1948/1972) *An outline of English phonetics.* Cambridge, U.K.: Cambridge University Press.

Jusczyk, P.W. (1997) *The discovery of spoken language.* Cambridge: Cambridge University Press.

Kelly, J. & Local, J.K. (1986) Long-domain resonance patterns in English. *Proceedings of the international conference on speech input/output.* Institute of Electronic Engineers, 304-8.

Kelly, J. & Local, J.K. (1989) *Doing phonology.* Manchester: University Press.

Kingston, J. & Diehl, R.L. (1995) Intermediate properties in the perception of distinctive feature values. In *Phonology and phonetic evidence: Papers in laboratory phonology IV* (B. Connell & A. Arvaniti, editors), pp. 7-27. Cambridge: Cambridge University Press.

Kwong, K. & Stevens, K.N. (1999) On the voiced-voiceless distinction for writer/rider. Research Laboratory of Electronics, Massachusetts Institute of Technology, *Speech Communication Group Working Papers* **XI**, 1-20.

Lisker, L. (1986) "Voicing" in English: a catalogue of acoustic features signaling /b/ versus /p/ in trochees. *Language & Speech* **29**, 3-11.

Löfquist, A. & Gracco, V. (1994) Tongue body kinematics in velar stop production: influences of consonant voicing and vowel context, *Phonetica* **51**, 52-67.

Local, J. (submitted) Variable domains and variable relevance: Interpreting phonetic exponents. *Journal of Phonetics.*

Marslen-Wilson, W. & Warren, P. (1994) Levels of perceptual representation and process in lexical access: Words, phonemes, and features. *Psychological Review* **101**, 653-675.

Massaro, D.W. & Cohen, M.M. (1983) Consonant/vowel ratio: An improbable cue in speech. *Perception & Psychophysics* **33**, 502-505.

Moore, B. & Glasberg, B.R. (1987) Formulae describing frequency selectivity as a function of frequency and level, and their use in calculating excitation patterns, *Hearing Research* **28**, 209–225.

Newton, D.E. (1996) The nature of resonance in English: an investigation into lateral articulations, *York Papers in Linguistics* **17**, 167-190.

Nittrouer, S., Studdert-Kennedy, M., & McGowan, R.S. (1989) The emergence of phonetic segments: Evidence from the spectral structure of fricative-vowel syllables spoken by children and adults. *Journal of Speech & Hearing Research* **32**, 120-132.

Nguyen, N. & Hawkins, S. (1998) Syllable-onset acoustic properties associated with syllable-coda voicing. *Proceedings of the 5th international conference on spoken language processing.* On CD-ROM: ICSLP-98. Paper #539.

Nguyen, N. & Hawkins, S. (1999) Implications for word recognition of phonetic dependencies between syllable onsets and codas. In *Proceedings of the XIVth international congress of phonetic sciences* (J.J. Ohala, Y. Hasegawa, M. Ohala, D. Granville & A.C. Bailey, editors), **1**, 647-650.

Ogden, R., Hawkins, S., House, J., Huckvale, M., Local, J., Carter, P., Dankovicová, J., & Heid, S. (2000). ProSynth: An integrated prosodic approach to device-independent, natural-sounding speech synthesis. *Computer Speech & Language* **14**(3), 177-210.

Peterson, G.E. & Lehiste, I. (1960) Duration of syllable nuclei in English. *Journal of the Acoustical Society of America* **32**, 693-703.

Pisoni, D.B (1991) Comment [on Bregman's paper]: Models of processing speech and nonspeech signals. In *Modularity and the motor theory of speech perception* (I.G. Mattingly & M. Studdert-Kennedy, editors), pp. 225-238. Hillsdale, NJ: Lawrence Erlbaum Associates.

Pisoni, D.B. & Lively, S.E. (1995) Variability and invariance in speech perception: A new look at some perceptual problems in perceptual learning. In *Speech perception and linguistic experience: Issues in cross-language research* (W. Strange, editor), pp. 411-432. Baltimore: York Press.

Plaut, D.C. & Kello, C.T. (1999) The emergence of phonology from the interplay of speech comprehension and production: A distributed connectionist approach. In *The emergence of language* (B. MacWhinney, editor), pp.381-415. Mahwah, NJ: Lawrence Erlbaum Associates.

Protopapas, A. (1999) Connectionist modeling of speech perception. *Psychological Bulletin* **125**, 410-436.

Raphael, L. (1972) Preceding vowel duration as a cue to the perception of the voicing characteristic of word-final consonants in American English. *Journal of the Acoustical Society of America* **51**, 1296-1303.

Remez, R.E. (2001) The interchange of phonology and perception considered from the perspective of organization. In *The role of speech perception phenomena in phonology* (E. V. Hume & K. A. Johnson, editors), pp. 27-52. New York: Academic Press.

Remez, R.E., Rubin, P.E., Berns, S.M., Pardo, J.S., & Lang, J.M. (1994) On the perceptual organization of speech. *Psychological Review* **101**, 129-156.

Slater, A. & Coleman, J.S. (1996) Non-segmental analysis and synthesis based on a speech database, *ICSLP 96*, **4**, 2379-2382.

Smith, R. (2001) Does memory for individual talkers help word segmentation? *TIPS: Temporal Integration in the Perception of Speech* (S. Hawkins & N. Nguyen, editors), p. 49. ISCA.

Sproat, R. & Fujimura, O. (1993) Allophonic variation in English /l/ and its implications for phonetic implementation. *Journal of Phonetics* **21**, 291-311.

Stevens, K.N. (1983) Design features of speech sound systems. In *The production of speech* (P. F. MacNeilage, editor), pp. 247-261. Berlin: Springer-Verlag.

Stevens, K.N. (1998) *Acoustic phonetics.* Cambridge, MA: MIT Press.

Streeter, L.A., & Nigro, G.N. (1979) The role of medial consonant transitions in word perception. *Journal of the Acoustical Society of America* **65**, 1533-1541.

Summers, W.V. (1987) Effects of stress and final-consonant voicing on vowel production: Articulatory and acoustic analyses. *Journal of the Acoustical Society of America* **82**, 847-863.

Summers, W.V. (1988) F1 structure provides information for final-consonant voicing. *Journal of the Acoustical Society of America* **84**, 485-492.

Suomi, K. (1993) An outline of a developmental model of adult phonological organization and behaviour. *Journal of Phonetics* **21**, 29-60.

Tunley, A. (1999) *Coarticulatory influences of liquids on vowels in English.* PhD dissertation, University of Cambridge, U.K.

van Santen, J.P.H., Coleman, J.S. & Randolph, M.A. (1992) Effects of postvocalic voicing on the time course of vowels and diphthongs. *Journal of the Acoustical Society of America* **92**(4/2), 2444.

Warren, P. & Marslen-Wilson, W. (1987) Continuous uptake of acoustic cues in spoken word recognition, *Perception & Psychophysics* **41**, 262-275.

Weismer, G. (1979) Sensitivity of voice-onset time (VOT) measures to certain segmental features in speech production. *Journal of Phonetics* **7**, 197-204.

West, P. (1999) The extent of coarticulation of English liquids: An acoustic and articulatory study. In *Proceedings of the XIVth International Congress of Phonetic Sciences* (J.J. Ohala, Y. Hasegawa, M. Ohala, D. Granville & A.C. Bailey, editors), **2**, 1901-1904.

West, P. (2000) Perception of distributed coarticulatory properties of English /l/ and /ɹ/. *Journal of Phonetics* **27**, 405-426.

Westbury, J. (1983) Enlargement of the supraglottal cavity and its relation to stop consonant voicing. *Journal of the Acoustical Society of America* **73**, 1322-1336.

Whalen, D.H. (1989) Vowel and consonant judgments are not independent when cued by the same information. *Perception & Psychophysics* **46**, 284-292.

Whalen, D.H., Best, C.T. & Irwin, J.R. (1997) Lexical effects in the perception and production of American English /p/ allophones. *Journal of Phonetics* **25**, 501-528.

Wolf, C.G. (1978) Voicing cues in English final stops. *Journal of Phonetics* **6**, 299-309.

Wright, R. (in press) Factors of lexical competition in vowel articulation. In *Papers in laboratory phonology VI* (J.K. Local, R.A. Ogden & R.A.M. Temple, editors). Cambridge University Press.

Zheng *et al.* COG freeware available from: ftp://ling.ohio-state.edu/pub/phonetics/specmom/

TABLE I. Mean duration (ms) of onset /l/ in voiced-coda and voiceless-coda monosyllables, for each S. The rightmost column (diff.) shows differences between contexts (voiced − voiceless coda).

speaker	voiced coda	voiceless coda	diff.
1	90.3	85.9	+4.4
2	78.8	74.6	+4.2
3	78.4	74.2	+4.2
4	86.8	83.0	+3.8

TABLE II. Mean duration (ms) of vowel in voiced-coda and voiceless-coda monosyllables, for each Speaker. The fourth column (diff.) shows differences between contexts (voiced − voiceless coda). The rightmost column (vl:v) shows the ratios of the vowel durations (voiceless-to-voiced coda context).

speaker	voiced coda	voiceless coda	diff.	vl:v
1	228	179	+49	0.79
2	159	121	+38	0.76
3	165	138	+27	0.84
4	181	138	+43	0.76

TABLE III. Mean frequency of F_2 (Hz) at the mid-point of onset /l/ in voiced-coda and voiceless-coda monosyllables. The rightmost column (diff.) shows differences between contexts (voiced – voiceless coda).

speaker	voiced coda	voiceless coda	diff.
1	1534	1536	-2
2	1292	1306	-14
3	1002	1033	-31
4	1112	1108	+4

TABLE IV. F ratios and significance levels (*p*) associated with significant changes in F_1 or F_2 frequency before a voiced coda relative to a voiceless coda. Columns headed 'voiced' and 'voiceless' show the frequencies (Hz) of the formant named in column 2 of the same row, for syllables with voiced and voiceless codas respectively; the column headed 'diff.' shows differences (Hz) between these two values (voiced – voiceless coda).

Speaker(s)	Formant	location	voiced	voiceless	diff.	F(1,38)	*p*
S1	F1	V-on	588	596	-8	7.089	.011
S1	F1	V-mid	575	594	-19	25.650	.000
S2	F1	l-mid	300	307	-7	8.039	.007
S2	F2	l-mid	1292	1306	-14	6.543	.015
S2	F1	V-on	534	541	-7	4.551	.039
S2	F1	V-mid	502	526	-24	46.425	.000
S3	F2	l-mid	1002	1033	-31	37.733	.000
S3	F1	V-on	424	437	-13	10.398	.003
S3	F2	V-on	1246	1257	-11	8.866	.005
S3	F1	V-mid	421	435	-14	10.278	.003
S4	F1	l-mid	325	339	-14	4.844	.034
S4	F1	V-on	540	553	-13	7.386	.010
S4	F2	V-on	1358	1348	+10	4.120	.049
S4	F1	V-mid	528	544	-16	5.613	.023
S1-S4	F1	l-mid	327	335	-8	6.674	.014
S1-S4	F2	l-mid	1235	1246	-11	16.092	.000
S1-S4	F1	V-on	521	532	-11	17.922	.000
S1-S4	F1	V-mid	507	525	-18	36.953	.000

TABLE V. F ratios and significance levels (*p*) associated with significant changes in the spectral center of gravity before a voiced coda relative to a voiceless coda. Columns headed ‘voiced’ and ‘voiceless’ show the mean center of gravity (Hz) at the location named in column 2, for syllables with voiced and voiceless codas respectively; the column headed ‘diff.’ shows differences between these two values (voiced – voiceless coda).

Speaker(s)	location	voiced	voiceless	diff.	F(1,38)	*p*
S1	l-mid	998	1004	-6	4.549	.039
S1	V-mid	1104	1113	-9	11.100	.002
S1	V-end	1033	1055	-22	31.996	.000
S2	l-mid	1092	1105	-13	5.698	.022
S2	V-end	1055	1092	-37	27.898	.000
S3	l-mid	927	941	-14	5.510	.024
S3	V-mid	1052	1046	+6	5.382	.026
S4	V-on	1071	1067	+4	5.299	.027
S4	V-end	1040	1063	-23	13.405	.001
S1-S3	l-mid	1006	1017	-11	16.924	.000
S1-S3	V-end	1035	1057	-22	41.771	.000

TABLE VI. Average changes in fundamental frequency before a voiced coda relative to a voiceless coda at five time locations. Negative differences indicate f_0 was lower before a voiced coda. Asterisks indicate significance levels (* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$; (*) approaching $p < 0.05$).

Speaker	l-mid	l-end	V-on	V-mid	V-end
S1	-0.26	-0.18	0.72	-1.20*	18.49***
S2	2.89(*)	0.83	2.89***	0.56	21.03***
S3	0.13	-1.54**	-0.97	-1.12*	9.65***
S4	-0.64	0.28	3.07***	0.64	6.48

TABLE VII. F ratios and significance levels (*p*) associated with significant changes in fundamental frequency before a voiced coda relative to a voiceless coda.

Speaker	location	F(1,38)	*p*
S1	V-mid	5.693	0.022
S1	V-end	25.434	0.000
S2	l-mid	3.761	0.060
S2	V-on	38.202	0.000
S2	V-end	105.747	0.000
S3	l-end	9.355	0.004
S3	V-mid	4.502	0.040
S3	V-end	17.303	0.000
S4	V-on	23.622	0.000

Figure legends

Figure 1. Change in F_1 and F_2 frequencies at three locations in the syllable, the middle of the /l/ (l-mid), 40-ms after vowel onset (V-on), and the midpoint of the vowel (V-mid), for individual speakers (left column) and the average for all four speakers (righthand panel). Within each panel, the lower line represents F_1 and the upper line represents F_2. Each shaded rectangle shows the mean difference in F_1 or F_2 frequency between minimal pairs with voiced vs. voiceless codas e.g. *blob – blop*. Negative differences mean the formant frequency was lower when the coda was voiced. Black bars represent differences that were statistically significant in ANOVAs at $p < 0.05$ or better. Grey bars represent non-significant effects ($p > 0.05$).

Figure 2. Same as Figure 1 except that only data for S1 and S2 are shown, and only for high-front, low-front, and low-back/central vowel groups (the back-rounded vowel group is omitted). Shading indicates statistical significance in ANOVAs: black, $p < 0.05$ or better; grey, $p > 0.05$.

Figure 3. The difference (Hz) for each speaker between F_2-F_1 in voiced codas, and F_2-F_1 in voiceless codas, at three locations in the syllable, as indicated. Negative differences indicate that the articulation is relatively more backed before a voiced coda. Positive differences indicate that it is relatively more fronted before a voiced coda. Shading indicates statistical significance in ANOVAs: black, $p < 0.05$ or better; grey, $p > 0.05$.

Figure 4. Left panels: Mean differences (Hz) in the spectral COG between voiced-coda and voiceless-coda syllables at five locations in the syllable, for each subject. Right panel: mean COG for Subjects 1-3. Locations (of the centre of the range over which the COG was calculated) are as follows: (1) the midpoint of the periodic part of /l/; (2) 24.5 ms before /l/ offset (so the right hand edge of the window fell at the segmentation point); (3) 40 ms after /l/ offset (roughly at the end of the /l/-to-V transition); (4) the midpoint of the vowel; (5) 30 ms before the end of the vowel. Negative differences mean the COG was lower when the coda was voiced. Shading indicates statistical significance in ANOVAs: black, $p < 0.05$ or better; grey, $p > 0.05$. See text for further details.

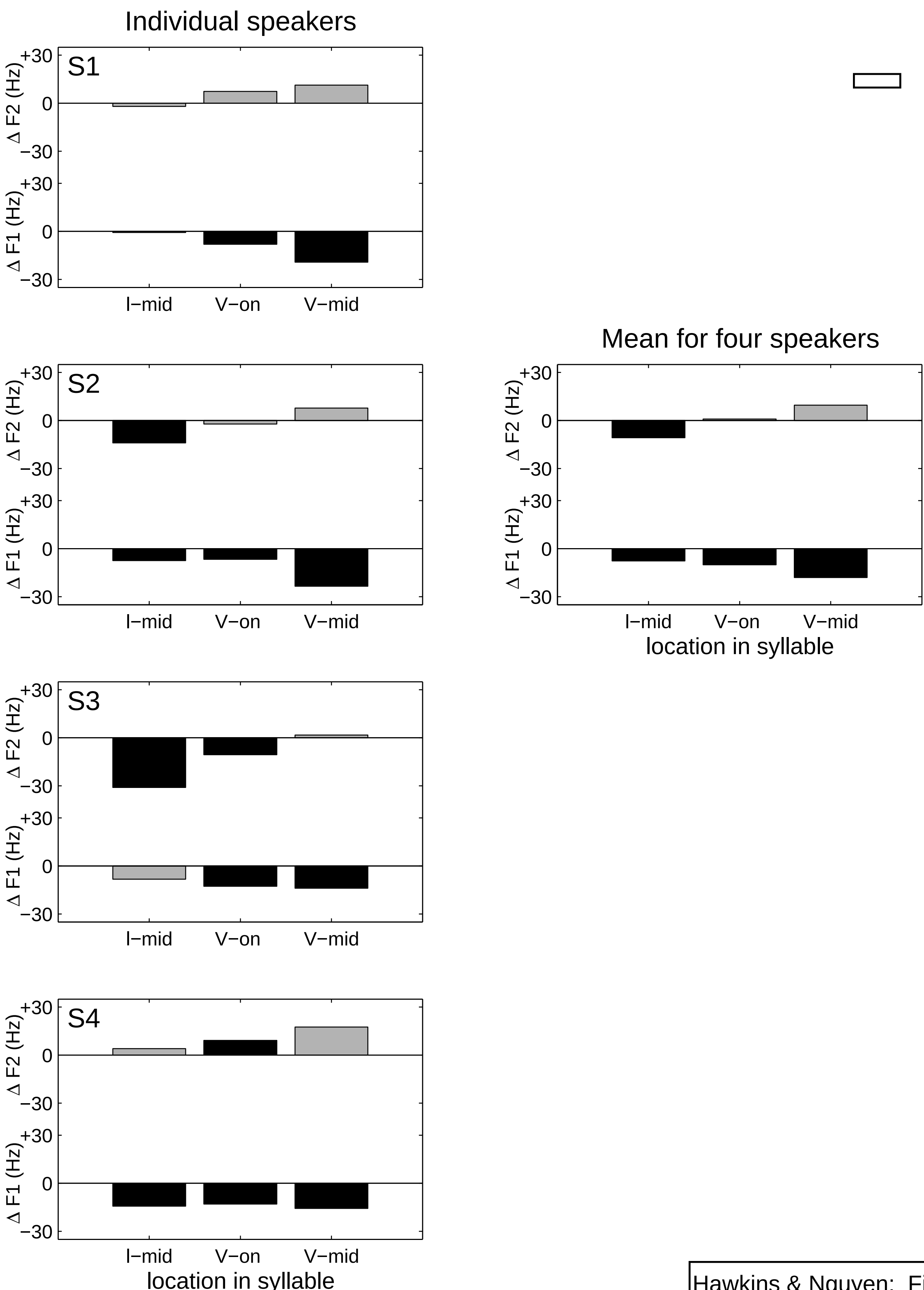
Figure 5. Top: spectrograms of *led* and *let* spoken by Speaker 1. Properties of the onset and nucleus of these syllables that can contribute to the distinction between voiced and voiceless codas are indicated. a = vowel

duration; b = F_1 vowel offset transition; c = F_1 vowel onset transition; 148 ms = duration of /l/ segment of *led*; 125 ms = duration of /l/ segment of *let*; F_2 = the second formant. Lower panel: lpc spectra centred at the midpoint of the /l/ segment in each word, showing the first three formant frequencies. The main difference is in the frequency of F_2, which is lower in *led* (heavy line).

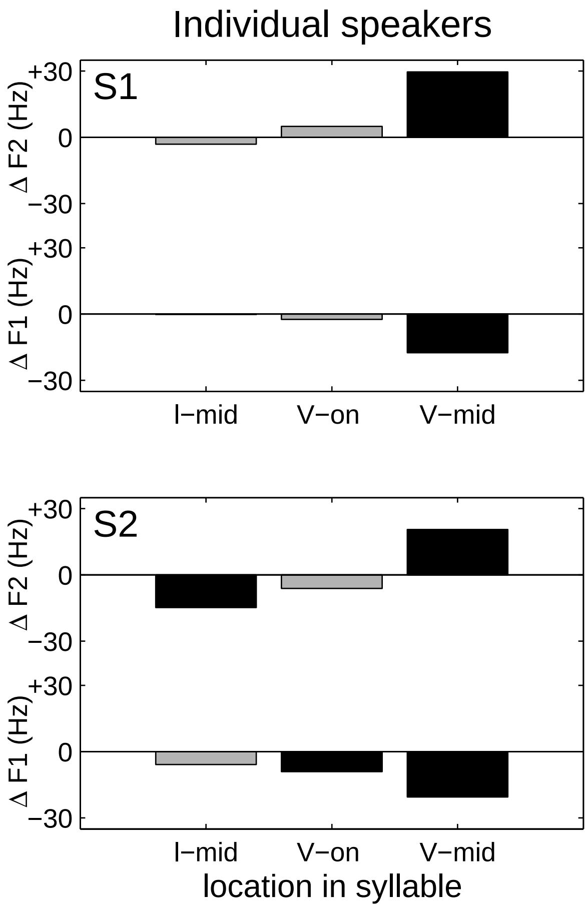
Footnotes

[1] Of the 22 other vowel x coda-voicing interactions, only three were significant: F1 at vowel onset for S1 ($F(3,35) = 3.249$, $p = 0.033$); F_1 at vowel mid-point for S3 ($F(3,35) = 3.559$, $p = 0.024$); and F_2 at vowel onset for S4, ($F(3,35) = 4.205$, $p = 0.012$). These interactions were due to the effect of coda voicing being restricted to the back-rounded vowel group for S1, and to the back-rounded and the low-back/central vowels for S3 and S4. (Before voiced codas, F_1 was lower for back-rounded vowels at vowel onset for S1 ($F(1,11) = 8.986$, $p = 0.012$) and at vowel midpoint for S3 ($F(1,11) = 13.998$, $p = 0.003$; also for low-back/central vowels at V-mid ($F(1,11) = 6.973$, $p = 0.027$)). For S4, F_2 at vowel onset was higher before voiced codas for back-rounded and low-back/central vowels ($F(1,11) = 6.925$, $p = .023$, $F(1,11) = 27.975$, $p = 0.001$ respectively). It is not clear to us how to interpret these patterns; however, the difference between the back-rounded group and the other groups for S1 and S2 are discussed below.

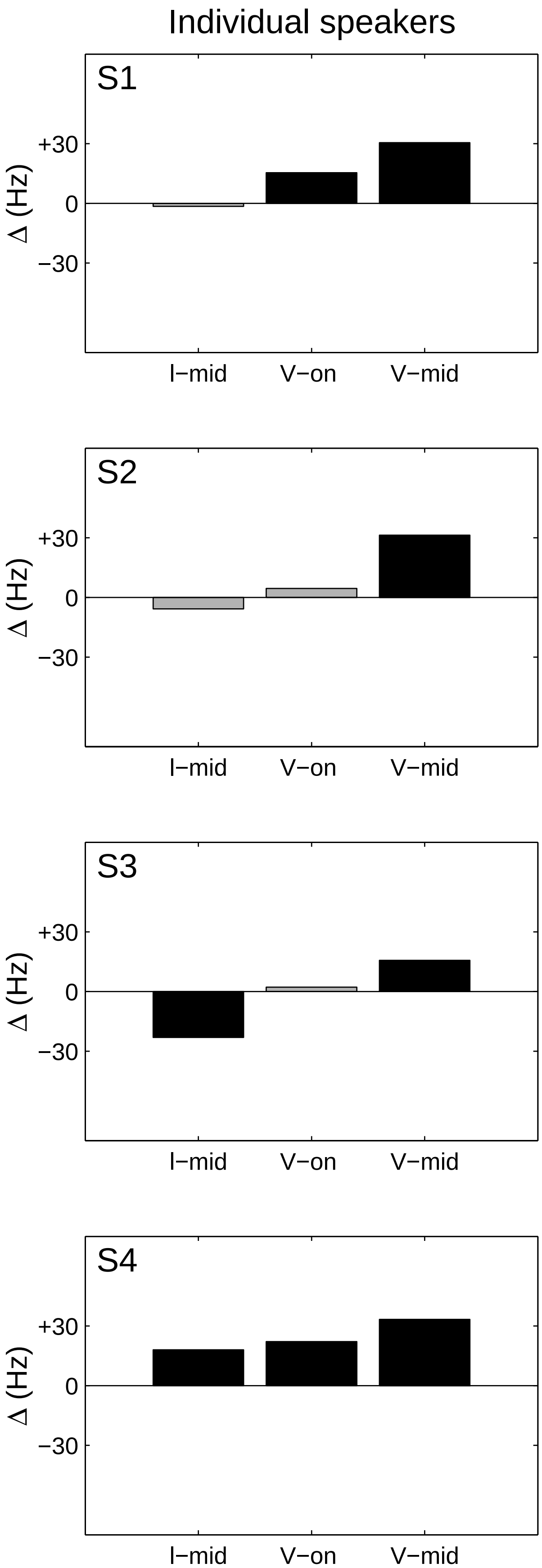
[2] These observations are broadly supported by ANOVAs on the degree of change in F_1 and F_2 frequencies between V-on and V-mid in voiced- compared with voiceless-coda syllables: for non-rounded vowels, and excluding S4, both F_1 and F_2 generally change more in the first half of the vowel when the coda is voiced. Details are available from the authors. These analyses are not reported here because they are extensive and add little to the point we are making. More work is needed to assess the details of the claim, aimed especially at distinguishing the influences of overall syllable duration, initial vowel transition duration, and rates of formant frequency change, as well as of vowel quality.



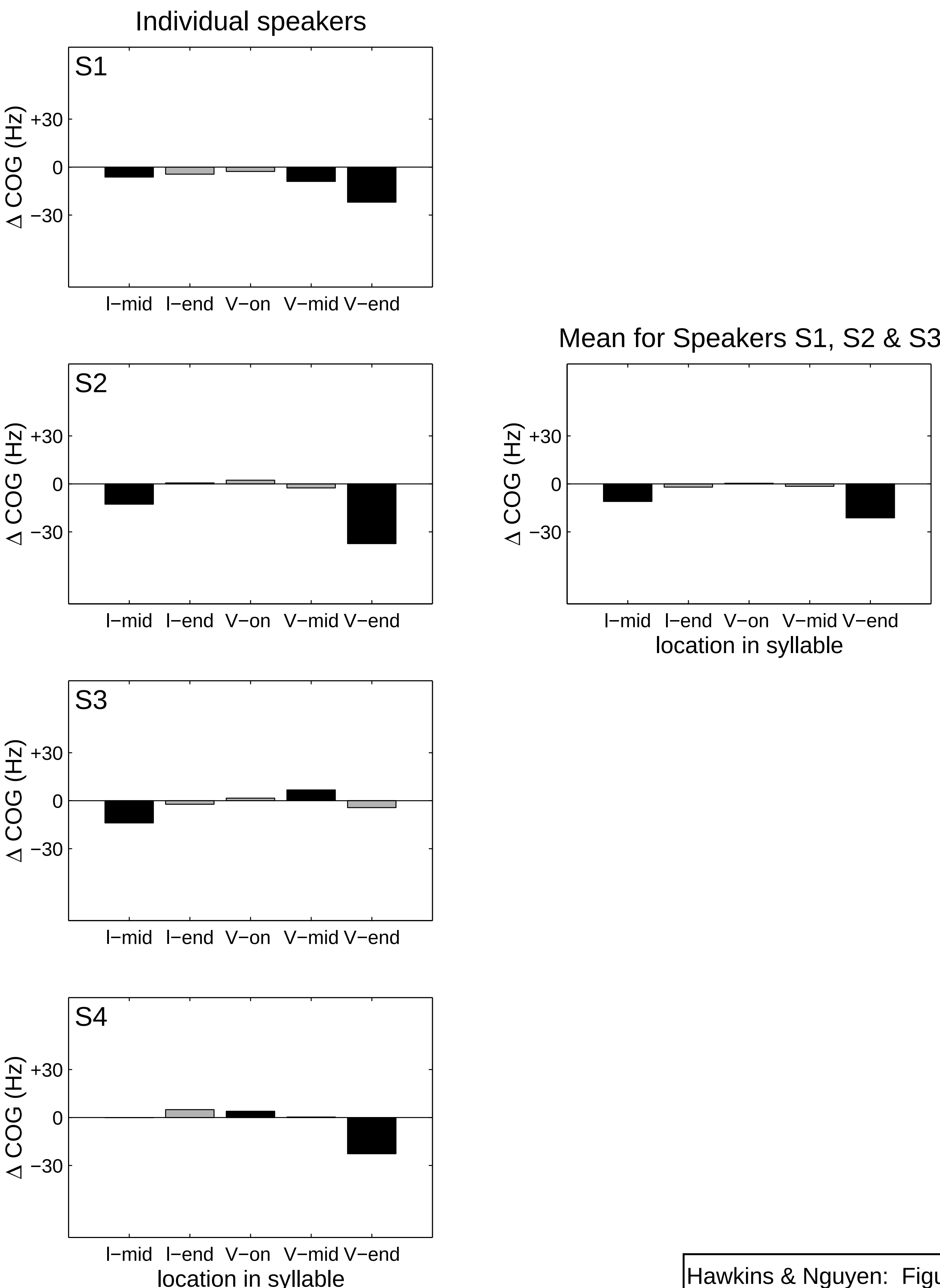
Hawkins & Nguyen: Figure 1



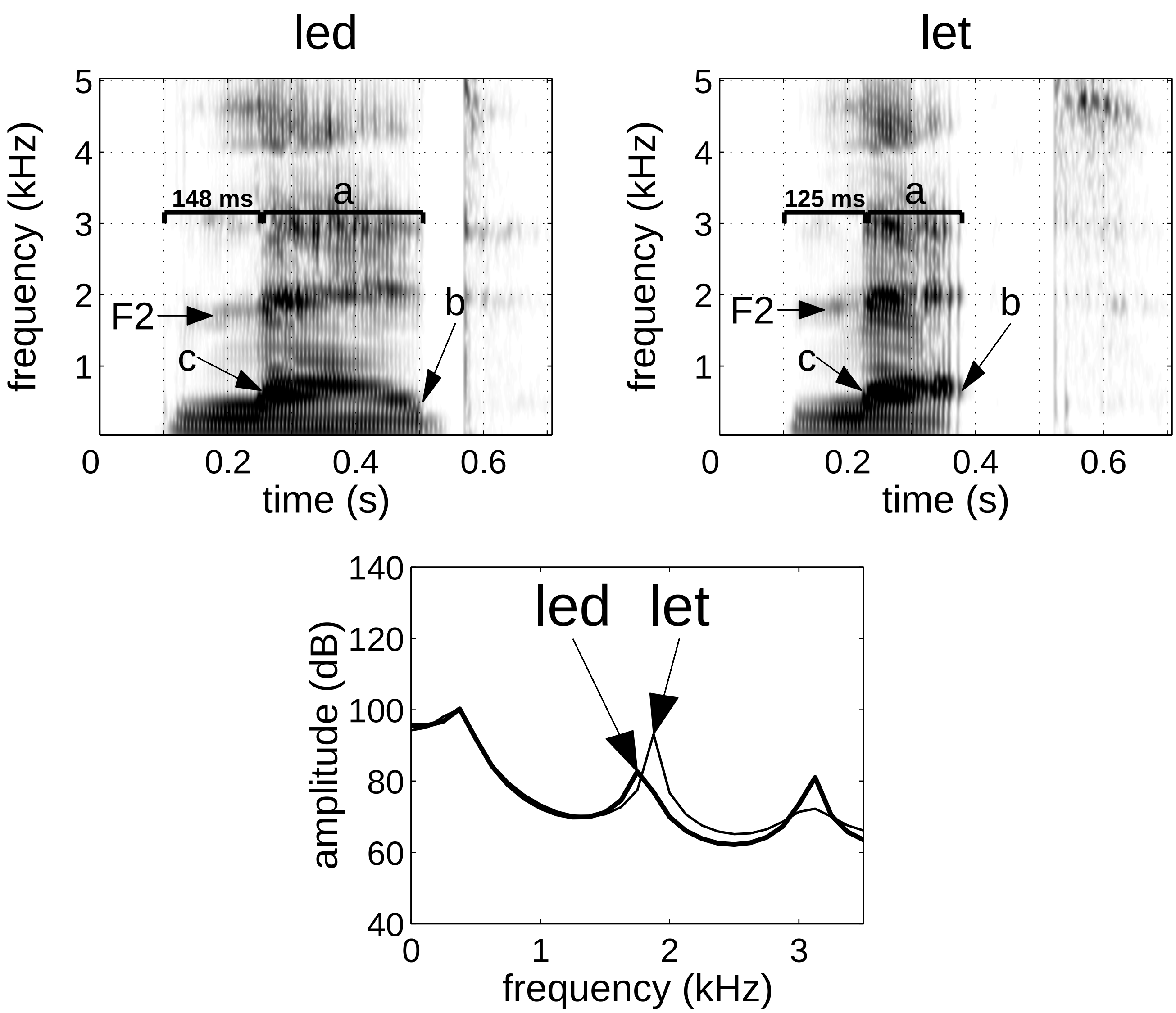
Hawkins & Nguyen: Figure 2



Hawkins & Nguyen: Figure 3



Hawkins & Nguyen: Figure 4



Hawkins & Nguyen: Figure 5